

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 28, 2002, 22:40:30 ; Search time 304.605 Seconds
(without alignments)
11355.910 Million cell updates/sec

Title: US-09-963-333-6

Perfect score: 1536

Sequence: 1 ggggggggggggaccacttg.....ataataaagaagtgtctgc 1536

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched:2185239 seqs, 1125999159 residues

Word size: 15

Total number of hits satisfying chosen parameters: 4914

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database:

N_Geneseq_101002.*
1: /SID82/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SID82/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SID82/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SID82/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SID82/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SID82/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SID82/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SID82/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SID82/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SID82/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SID82/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SID82/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SID82/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SID82/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SID82/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SID82/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SID82/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SID82/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SID82/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SID82/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SID82/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SID82/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SID82/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SID82/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1394	90.8	1536	ABK43335	Human Thymidylate
2	1346	87.6	3298	AAS94945	Human DNA sequence
3	536	34.9	1539	AAS84960	DNA encoding novel
4	532	34.6	535	ABL38559	Human colon tumour
5	487	31.7	18596	AAF31109	Thymidylate synthase
6	487	31.7	18596	22 AAC91215	Human thymidylate
7	487	31.7	18596	22 AAB95092	Gene #1590 used to
8	487	31.7	18596	24 ABK43334	Human thymidylate
9	487	31.7	18596	24 ABL62854	Breast cancer rela

10	487	31.7	18596	24	ABL63078	Breast cancer rela
11	487	31.7	18596	24	ABL67927	Ovary cancer relat
12	487	31.7	45716	24	ABA93401	Human rts-alpha ge
13	487	31.7	45989	24	ABA93402	Human rts-beta gen
14	379	24.7	566	20	AA24270	Human thymidylate
15	178	11.6	651	23	AAS84956	DNA encoding novel
16	168	10.9	1817	24	ABA93399	Human rts-alpha en
17	168	10.9	1857	22	AA159987	Human polynucleoti
18	150	9.8	196	22	AAS04589	Gene expression pr
19	142	9.2	346	21	AA78156	CDNA encoding huma
20	142	9.2	346	22	AA128894	Colon tumour relat
21	107	7.0	1131	24	ABK43330	Human GNEHK genomic
22	107	7.0	1161	24	ABK43288	Human CDNA encodin
23	79	5.1	118	19	AA10927	Human biallelic po
24	77	5.0	1692	22	ABA09660	Human bone marrow
25	77	5.0	3422	23	AAS84959	DNA encoding novel
26	60	3.9	196	24	AAB33855	Human spliced tran
27	59	3.8	2945	23	AAS84962	Gene expression pr
28	40	2.6	2945	23	AAS84962	DNA encoding novel
29	30	2.0	30	20	AA24268	Thymidylate synthase
30	30	2.0	30	20	AA24269	Thymidylate synthase
31	30	2.0	30	20	AAV71964	Thymidylate synthase
32	30	2.0	30	20	AAV71965	Thymidylate synthase
33	30	2.0	36	20	AA24274	Thymidylate synthase
34	30	2.0	36	20	AA24271	Thymidylate synthase
35	30	2.0	36	20	AA24272	Thymidylate synthase
36	30	2.0	36	20	AA24273	Thymidylate synthase
37	30	2.0	230	23	ABV57305	Human prostate exp
38	29	1.9	29	22	AAH41174	Human thymidylate
39	27	1.8	764	24	ABQ39528	Oligonucleotide fo
40	27	1.8	764	24	ABQ39529	Oligonucleotide fo
41	27	1.8	1186	24	ABK40090	Human chemically p
42	26	1.7	26	19	AA09473	Human biallelic po
43	26	1.7	26	19	AAV52596	Primer hts-3a, use
44	26	1.7	128139	24	AA164291	RRV genome nucleot
45	26	1.7	133719	21	AA64754	Macaca mulatta rha

ALIGNMENTS

RESULT 1

ABK43335 ID ABK43335 standard; cDNA; 1536 BP.

XX AC ABK43335;

XX DT 05-JUN-2002 (first entry)

XX XX Human Thymidylate synthase cDNA sequence.

XX DE HKNG1; ss; gene; chromosome 18p; bipolar affective disorder; BAD;

XX DE severe bipolar affective (mood) disorder; BP-I; schizophrenia; TS;

XX DE Hong Kong new gene 1; antimanic; antidepressant; neuroleptic;

XX DE Thymidylate synthase.

XX OS Homo sapiens.

XX XX WO200210366-A2.

XX PD 07-FEB-2002.

XX XX 02-AUG-2001; 2001WO-US24417.

XX XX 02-AUG-2000; 2000US-0631275.

XX XX 28-NOV-2000; 2000US-0725544.

XX PA (MILL-) MILLENNIUM PHARM INC.

XX PA (REGC) UNIV CALIFORNIA.

XX PI Chen H, Freimer NB, Novak T;

XX DR WPI; 2002-195962/25.

Query Match 68.1%; Score 1045.4; DB 9; Length 1077;
Best Local Similarity 97.9%; Pred. No. 1.7e-261;
Matches 1049; Conservative 5; Mismatches 18; Indels 0; Gaps 0;

Qy 438 CAATGATCCCGAGACTTTTGGACAGCTGGATTTCCACACAGAGAGAGAGGGGACTT 497
Db 1072 CCAATGATCCCGAGACTTTTGGACAGCTTTGGATTTCCACACAGAGAGAGGGGACTT 1013

Qy 498 GGGCCAGTTTATGGCTTCAGTGAGGCAATTTGGGGCAGATACAGAGATATGGAATC 557
Db 1012 GGGCCAGTTTATGGCTTCAGTGAGGCAATTTGGGGCAGATACAGAGATATGGAATC 953

Qy 558 AGATTATTCAGGACAGGAGTTCACCACTGCAAGAGTATTTGACACCATCAAAACCAA 617
Db 952 AGATTATTCAGGACAGGAGTTCACCACTGCAAGAGTATTTGACACCATCAAAACCAA 893

Qy 618 CCCTGACGACAGAGAAATCATCATGTGGCTTGAATCCAAAGAGATTTCTCTCTGATGGC 677
Db 892 CCCTGACGACAGAGAAATCATCATGTGGCTTGAATCCAAAGAGATTTCTCTCTGATGGC 833

Qy 678 GCTGCTCCATGCTCCCTCTGCGAGTCTTATGTGGTGAACAGTGAGCTGCTCTGCCA 737
Db 832 GCTGCTCCATGCTCCCTCTGCGAGTCTTATGTGGTGAACAGTGAGCTGCTCTGCCA 773

Qy 738 GCTGTACAGAGATCGGAGACATGGGCTCGTGTGCTTTCAACATCGCCAGCTAGC 797
Db 772 GCTGTACAGAGATCGGAGACATGGGCTCGTGTGCTTTCAACATCGCCAGCTAGC 713

Qy 798 CTTGCTCAGCTACATGATTTGCCACATCACGGGCTCAAGCAGGAGTATTTATACACAC 857
Db 712 CTTGCTCAGCTACATGATTTGCCACATCACGGGCTCAAGCAGGAGTATTTATACACAC 653

Qy 858 TTTCGGAGATGACATATTTACTGATACATCGAGCTCAAAATTCAGCTTCAGCG 917
Db 652 TTTCGGAGATGACATATTTACTGATACATCGAGCTCAAAATTCAGCTTCAGCG 593

Qy 918 AGAACCCAGACCTTTCCCAAGCTCAGAGTCTTTCGAAAAGTTGAGAAATTTGATGACT 977
Db 592 AGAACCCAGACCTTTCCCAAGCTCAGAGTCTTTCGAAAAGTTGAGAAATTTGATGACT 533

Qy 978 CAAAGCTGAAGACTTTTCAAGTGAAGGTCACATCCGCTCCAACTATTTAAATGGAAT 1037
Db 532 CAAAGCTGAAGACTTTTCAAGTGAAGGTCACATCCGCTCCAACTATTTAAATGGAAT 473

Qy 1038 GCCTGTATGAGGTCCTTCAAGAGCTGAGATTTCTGAGTCTTTAGGGTGGC 1097
Db 472 GCCTGTATGAGGTCCTTCAAGAGCTGAGATTTCTGAGTCTTTAGGGTGGC 413

Qy 1098 TGGATGCGGAGTAAAGTCTTTTTCCTTAAAGAAAGAAAGAACTAGCTCAAAATCT 1157
Db 412 TGGATGCGGAGTAAAGTCTTTTTCCTTAAAGAAAGAAAGAACTAGCTCAAAATCT 353

Qy 1158 GTCCGTGACCTATCAGTTATTAATTTTAAAGATTTGACACTGGCAATGTAAGTGGC 1217
Db 352 GTCCGTGACCTATCAGTTATTAATTTTAAAGATTTGACACTGGCAATGTAAGTGGC 293

Qy 1218 CAGTTCTTTCCATAATAAAGGCTTTGATTAACCTCACTCAGGGTATCTGACAAATCTGA 1277
Db 292 CAGTTCTTTCCATAATAAAGGCTTTGATTAACCTCACTCAGGGTATCTGACAAATCTGA 233

Qy 1278 GGTATGAACAAAGTGGAGAGATGAATATGATGCTCTTACAAAAACATGATATGTC 1337
Db 232 GGTATGAACAAAGTGGAGAGATGAATATGATGCTCTTACAAAAACATGATATGTC 173

Qy 1338 ATTTCAATCCCACTATTAATAAGAGGTTGGTGAATTTCAAGCTATTTTTCGAATA 1397
Db 172 ATTTCAATCCCACTATTAATAAGAGGTTGGTGAATTTCAAGCTATTTTTCGAATA 113

Qy 1398 TTTTGAATAATTTTGAATTTTCAAGACTATTTCCCTCAAAATCTGAGGAGCTGAGTAA 1457
Db 112 TTTTGAATAATTTTGAATTTTCAAGACTATTTCCCTCAAAATCTGAGGAGCTGAGTAA 53

Qy 1458 CACCATCGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1509

Db 52 CACCATGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1

RESULT 4
AL547612 1020 bp mRNA linear EST 16-FEB-2001
LOCUS AL547612 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0D1008YH21 5
DEFINITION prime, mRNA Sequence.
ACCESSION AL547612
VERSION AL547612.1 GI:12881832
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1. (bases 1 to 1020)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
Location/Qualifiers
source
1..1020
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D1008YH21"
/clone_lib="LTI_NFL006_PL2"
/tissue_type="placenta"
/notes="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 237 a 291 c 276 g 215 t 1 others
ORIGIN

Query Match 66.3%; Score 1018; DB 9; Length 1020;
Best Local Similarity 99.8%; Pred. No. 2.4e-254;
Matches 1018; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 35 CGCGCGGCACTTCGCTGCTCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 94
Db 1 CGCGCGGCACTTCGCTGCTCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 60

Qy 95 CGCGCGGCGCATCCCTGTGGCGCGCTCGAGCTGCGCGCGCGCGCGCGCGCGCGCGCG 154
Db 61 CGCGCGGCGCATCCCTGTGGCGCGCTCGAGCTGCGCGCGCGCGCGCGCGCGCGCG 120

Qy 155 CACAGAGCGGAGACCGCGAGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 214
Db 121 CACAGAGCGGAGACCGCGAGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 180

Qy 215 TCCACACATCTCTGCTGGCGGTGAGAGAGACCGCGCGCGCGCGCGCGCGCGCGCGCT 274
Db 181 TCCACACATCTCTGCTGGCGGTGAGAGAGACCGCGCGCGCGCGCGCGCGCGCGCT 240

Qy 275 CGGTATTCGATCGAGCGCGCTACAGCTCAGAGATGAATTCCTCTGCTGACAAACA 334
Db 241 CGGTATTCGATCGAGCGCGCTACAGCTCAGAGATGAATTCCTCTGCTGACAAACA 300

Qy 335 AACGTGTGTTTGGAAAGGTTTGGAGAGTTTCTGCTGCTGCTTATCAAGGGATCCACAA 394
Db 301 AACGTGTGTTTGGAAAGGTTTGGAGAGTTTCTGCTGCTGCTTATCAAGGGATCCACAA 360

Qy 395 ATGCTTAAGAGCTGCTCTCCAAAGGAGTGAATTCGGGATGCCAATGGATCCCGAGACT 454

Shevchenko, V., Wetherby, K.D., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks S., Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karling, S., Legaspi, R., Lim, M., Maduro, Q.L., Masiello, C., Mastrian, S.B., McLoakes, J.C., McDowell, J., Pearson, R., Snyder, B., Stantirispob, S., Thomas, P.J., Tsang, L.H., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A., Zhang, L.H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 4 Row: h Column: 10
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA g1: 10835001
This clone has the following problem: no 5' EST match.

FEATURES
source

```
1. 1533
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3141419"
/tissue type="Ovary, ade
/clone lib="NIH MGC 9"
/lab_host="DH10B-R"
/notes="vector. n0TB7"
```

BASE COUNT	424 a	354 c	382 g	373 t
ORIGIN				

Query Match 96.3%; Score 1479.4; DB 11; Length 1533;
Best Local Similarity 99.7%; Pred. No. 0;
0; Mismatches 4; Indels 0; Gaps 0;

53 GGCCTCGTCGCGCGGCGCACTTCGGCTGGCTCCGTCGCCCGCGCGCATGCTG 112

Db 9 G C C T C G G T C C C G C G C G C C A C T T C G C C T G C C T C C G T C C C C G C C G C G C C A T G C C T G 68

QY 113 TGGCCGGCTGGAGCTGCCGGCGCCGGCCCTTGTCCGCCCGCCGACAGGAGCGGAGCCCG 172

Db 69 TGGCCGGCTCGGAGTGCCTGGCGCGCGGCCCTTGGCCCCCCCCGACAGGAGGGGACGCCG 128

QY 173 AGCGCGTCCGCCGACCGCGGAGCTGCAGTACCTGGGGCAGATCCAAACACATCTTCCGT 232

D_b 129 AGCGCGTCCGCCCGCACCGGGAGCTGCAGTACCTGGGGCAGATCCACACATCC100G1 188

QY
233 GCGGCGTCAGGAAGGACGCACGGCACCGGCCCTGTCCGATATCGGCATCGACG

D6
189 GUGGUGTCAGGAAGGACCGCAGGGGCACCCGCCCTGTCCCATTCCCC

[illegible][illegible][illegible]

500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000

2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525
 526

Q-1 473 TCTCTCA CCGACACACAGACGGGACCTTTTATGGCTTCCAGTGGAGGCATTTTTG 532

429 TCTCCACGAGAGAGAGGGGCACTTGGGCCCCAGTTTATGGCTTCCAGTGGAGGCATTG 488

533 GGGCAGAATACAGAGATATGGAAATCAGATTATTCAGGACAGGAGCTTGACCAACTGCAAA 592

Db 489 GGGCAGAATACAGAGATATGGAAATCAGATTATTCAGGACAGGAGTTGACCAACTGCCAA 548

QY 593 GAGTGATTGACACCATCAAACCAACCTTGACGACAGAAGATCATCATGTGGCGTTGGA 652

Db 549 GAGTGAATTGACACCATCAAAACCAACCTGACGACAGAAGATCATCATGTGGCTTGA 608

QY 653 ATCCAGAGATCTTCTCTGATGGGCTGCCTCCATGCCATGCCCTCTGCCAGTTCTATG 712

Db 609 ATCCAAGAGATCTTCTCTGATGGCGTGCCCTCCATGCCATGCCCTCTGCCAGTTCTATG 668

[illegible]

RESULT 2

AL575295/c

DEFINITION

ACCESSION
VERSION

KEYWORDS
SOURCE

ORGANISM

REFERENCE

AUTHORS	TITLE
...	...

Accession	Length	Type	Source	Accession	Length	Type	Source
AL575295	1089 bp	mRNA	linear	EST 16-FEB-2001			
U757295							
U757296							
U757297							
U757298							
U757299							
U757300							
U757301							
U757302							
U757303							
U757304							
U757305							
U757306							
U757307							
U757308							
U757309							
U757310							
U757311							
U757312							
U757313							
U757314							
U757315							
U757316							
U757317							
U757318							
U757319							
U757320							
U757321							
U757322							
U757323							
U757324							
U757325							
U757326							
U757327							
U757328							
U757329							
U757330							
U757331							
U757332							
U757333							
U757334							
U757335							
U757336							
U757337							
U757338							
U757339							
U757340							
U757341							
U757342							
U757343							
U757344							
U757345							
U757346							
U757347							
U757348							
U757349							
U757350							
U757351							
U757352							
U757353							
U757354							
U757355							
U757356							
U757357							
U757358							
U757359							
U757360							
U757361							
U757362							
U757363							
U757364							
U757365							
U757366							

prime, mRNA sequence.

AL575295
AL575295.1 GI:12936324

EST.
human.

Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Primates; Hominidae; Homo

Mammalia; Euteleia; Placentalia; Carnivora; Canidae; Felidae; Mustelidae; Ursidae; Procyonidae; Mephitidae; Viverridae; Hyaenidae; Perissodactyla; Artiodactyla; Cetartiodactyla; Chiroptera; Insectivora; Soricomorpha; Rodentia; Lagomorpha; Monotremata; Marsupialia; Eutheria; Primates; Scandentia; Nonhuman Primates; Hominoidea; Hominidae; Homo

1 (bases 1 to 1089)

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization

Unpublished (2001)
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
 Location/Qualifiers
 1. .1089
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CSODI061YF09"
 /clone_lib="LTI_NFL006_PL2"
 /tissue_type="placenta"
 /note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 298 a 253 c 217 g 318 t 3 others

ORIGIN

Query Match 68.2%; Score 1048.2; DB 9; Length 1089;
 Best local Similarity 98.9%; Pred. No. 3.2e-262;
 Matches 1072; Conservative 2; Mismatches 8; Indels 2; Gaps 2;

QY 432 GGATGCAATGATGCCGAGACTTTTGGACAGCTGGGATTCACACAGAGAAGG 491
 Db 1083 GGATGCCAATGGAYCCGAGAC-TTTGGACAGCT-GGATTCACACAGAGA 1026

QY 492 GGACTTGGGCCAGTTATAGCTTCCAGTGAGGCATTTGGGCGAGATACAGAGATAT 551
 Db 1025 GGACTTGGGCCAGTTATAGCTTCCAGTGAGGCATTTTGGGCGAGATACAGAGATAT 966

QY 552 GGAATCAGATTATCAGGACGGAGTTGACCACTGCAAGAGTGATGACACATCAA 611
 Db 965 GGAATCAGATTATCAGGACGGAGTTGACCACTGCAAGAGTGATGACACATCAA 906

QY 612 AACCAACCTGACGACAGAGAAATCATCATGTGGCTTGGAAATCCAGAGATCTTCTCT 671
 Db 905 AACCAACCTGACGACAGAGAAATCATCATGTGGCTTGGAAATCCAGAGATCTTCTCT 846

QY 672 GATGGCGCTGCTCCATGCGCATGCGCTCTGCCAGTTCTATGTGTGCAACAGTGTGTC 731
 Db 845 GATGGCGCTGCTCCATGCGCATGCGCTCTGCCAGTTCTATGTGTGCAACAGTGTGTC 786

QY 732 CTGCCAGCTGTACAGAGATCGGAGACATGGGCGCTCGGTGTGCTTTCAACATCGCCAG 791
 Db 785 CTGCCAGCTGTACAGAGATCGGAGACATGGGCGCTCGGTGTGCTTTCAACATCGCCAG 726

QY 792 CTAGCGCTGCTCAGTACATGATTTGGGCGACATCAGGGCGCTGAGCCAGTGTCTTAT 851
 Db 725 STAGCGCTGCTCAGTACATGATTTGGGCGACATCAGGGCGCTGAGCCAGTGTCTTAT 666

QY 852 ACACATTTGGGAGATGCACATATTTACCTGAATCATCATCGAGCCACTGAAATTCAGCT 911
 Db 665 ACACATTTGGGAGATGCACATATTTACCTGAATCATCATCGAGCCACTGAAATTCAGCT 606

QY 912 TCAGCGAGAACCCAGACCTTTCCCAAGCTTCAGATTCTTCGAAAAGTTGAGAAATTTGA 971
 Db 605 TCAGCGAGAACCCAGACCTTTCCCAAGCTTCAGATTCTTCGAAAAGTTGAGAAATTTGA 546

QY 972 TGACTTCAGAGCTGGAAGACTTTCAGATTGAAGGTGTAATCCGCTCAACTATTAAAT 1031
 Db 545 TGACTTCAGAGCTGGAAGACTTTCAGATTGAAGGTGTAATCCGCTCAACTATTAAAT 486

QY 1032 GGAATAGCTGTTTGGGTGCTTTTCAAGGAGCTGAGGATATTTGTCAGTCTTTAGGGG 1091
 Db 485 GGAATAGCTGTTTGGGTGCTTTTCAAGGAGCTGAGGATATTTGTCAGTCTTTAGGGG 426

QY 1092 TTGGGCTGGATGCGAGGTAAAGTTCTTTTGTCTTAAAGAAAGAGAACTAGCTCAA 1151
 Db 425 TTGGGCTGGATGCGAGGTAAAGTTCTTTTGTCTTAAAGAAAGAGAACTAGCTCAA 366

QY 1152 AAATGTGTCGGTGAACCTATCAGTTATTAATTTTAAAGCATCTTCCCACTGGCAAAATGTA 1211
 Db 365 AAATGTGTCGGTGAACCTATCAGTTATTAATTTTAAAGCATCTTCCCACTGGCAAAATGTA 306

QY 1212 CTGTGCCAGTTCTTTCCATTAATAAAGGCTTTGAGTTAACTCACTGAGGGTATCTGACAA 1271
 Db 305 CTGTGCCAGTTCTTTCCATTAATAAAGGCTTTGAGTTAACTCACTGAGGGTATCTGACAA 246

QY 1272 TGCTCAGGTTATGACAAAGTGGAGAGATGAATGTATGTCTCTTAGCAAAAACATGT 1331
 Db 245 TGCTCAGGTTATGACAAAGTGGAGAGATGAATGTATGTCTCTTAGCAAAAACATGT 186

QY 1332 ATGTGCATTTCAATCCACAGTCTTATTAAGAAGGTTGGTGAATTTTCAAGCTATTTTT 1391
 Db 185 ATGTGCATTTCAATCCACAGTCTTATTAAGAAGGTTGGTGAATTTTCAAGCTATTTTT 126

QY 1392 GGAATATTTTGAATATTTTAAAGATTTTCAAGCTATTTCCCTCAAACTCGAGGAGCT 1451
 Db 125 GGAATATTTTGAATATTTTAAAGATTTTCAAGCTATTTCCCTCAAACTCGAGGAGCT 66

QY 1452 GAGTAACACCATCGATCATGATGTAGAGTGTGTTATCAAACTTTANAGTTGTTTATATG 1511
 Db 65 GAGTAACACCATCGATCATGATGTAGAGTGTGTTATCAAACTTTANAGTTGTTTATATG 6

QY 1512 TTGC 1515
 Db 5 TTGC 2

RESULT 3
 AL572496/c
 LOCUS
 DEFINITION AL572496 LTI_NFL006_PL2 Homo sapiens cDNA clone CSODI008YH21 3
 prime, mRNA sequence.
 ACCESSION AL572496
 VERSION AL572496.1 GI:12930822
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1077)
 AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
 Location/Qualifiers
 1. .1077
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CSODI008YH21"
 /clone_lib="LTI_NFL006_PL2"
 /tissue_type="placenta"
 /note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 294 a 251 c 214 g 309 t 9 others

ORIGIN

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

M nucleic - nucleic search, using sw model

un on: November 28, 2002, 19:35:09 ; Search time 2460.31 Seconds
(without alignments)
12130.226 Million cell updates/sec

itle: US-09-963-333-7

effect score: 1187

equences: 1 gatcgccactgactcca.....gagggagcgcgctgggga 1187

coring table: IDENTITY NUC

Gapop 10_0 , Gapext 1.0

earched: 24791104 seqs, 12571243825 residues

otal number of hits satisfying chosen parameters: 49582208

inimum DB seq length: 0

aximum DB seq length: 2000000000

ost-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

atabase : Pending Patents NA Main.*

1: /cgn2_6/ptodata/2/pna/US0600 COMB.seq.*
2: /cgn2_6/ptodata/2/pna/US06 COMB.seq.*
3: /cgn2_6/ptodata/2/pna/US07 COMB.seq.*
4: /cgn2_6/ptodata/2/pna/US08 COMB.seq.*
5: /cgn2_6/ptodata/2/pna/US08 COMB.seq.*
6: /cgn2_6/ptodata/2/pna/US08 COMB.seq.*
7: /cgn2_6/ptodata/2/pna/US08 COMB.seq.*
8: /cgn2_6/ptodata/2/pna/US08 COMB.seq.*
9: /cgn2_6/ptodata/2/pna/US08 COMB.seq.*
10: /cgn2_6/ptodata/2/pna/US08 COMB.seq.*
11: /cgn2_6/ptodata/2/pna/US08 COMB.seq.*
12: /cgn2_6/ptodata/2/pna/US08 COMB.seq.*
13: /cgn2_6/ptodata/2/pna/US08 COMB.seq.*
14: /cgn2_6/ptodata/2/pna/US09 COMB.seq.*
15: /cgn2_6/ptodata/2/pna/US09 COMB.seq.*
16: /cgn2_6/ptodata/2/pna/US09 COMB.seq.*
17: /cgn2_6/ptodata/2/pna/US09 COMB.seq.*
18: /cgn2_6/ptodata/2/pna/US09 COMB.seq.*
19: /cgn2_6/ptodata/2/pna/US09 COMB.seq.*
20: /cgn2_6/ptodata/2/pna/US09 COMB.seq.*
21: /cgn2_6/ptodata/2/pna/US09 COMB.seq.*
22: /cgn2_6/ptodata/2/pna/US09 COMB.seq.*
23: /cgn2_6/ptodata/2/pna/US09 COMB.seq.*
24: /cgn2_6/ptodata/2/pna/US09 COMB.seq.*
25: /cgn2_6/ptodata/2/pna/US09 COMB.seq.*
26: /cgn2_6/ptodata/2/pna/US09 COMB.seq.*
27: /cgn2_6/ptodata/2/pna/US09 COMB.seq.*
28: /cgn2_6/ptodata/2/pna/US09 COMB.seq.*
29: /cgn2_6/ptodata/2/pna/US09 COMB.seq.*
30: /cgn2_6/ptodata/2/pna/US09 COMB.seq.*
31: /cgn2_6/ptodata/2/pna/US09 COMB.seq.*
32: /cgn2_6/ptodata/2/pna/US09 COMB.seq.*
33: /cgn2_6/ptodata/2/pna/US09 COMB.seq.*
34: /cgn2_6/ptodata/2/pna/US09 COMB.seq.*
35: /cgn2_6/ptodata/2/pna/US09 COMB.seq.*
36: /cgn2_6/ptodata/2/pna/US09 COMB.seq.*
37: /cgn2_6/ptodata/2/pna/US09 COMB.seq.*
38: /cgn2_6/ptodata/2/pna/US09 COMB.seq.*
39: /cgn2_6/ptodata/2/pna/US09 COMB.seq.*
40: /cgn2_6/ptodata/2/pna/US09 COMB.seq.*
41: /cgn2_6/ptodata/2/pna/US09 COMB.seq.*
42: /cgn2_6/ptodata/2/pna/US09 COMB.seq.*
43: /cgn2_6/ptodata/2/pna/US09 COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1178	99.2	1187	25	US-09-658-659-7	Sequence 7, Appli
2	1178	99.2	1187	36	US-09-963-665-7	Sequence 7, Appli
3	1178	99.2	1187	36	US-09-963-677-7	Sequence 7, Appli
4	1178	99.2	1187	36	US-09-963-333-7	Sequence 7, Appli
5	1167	98.3	18596	22	US-09-577-266-11	Sequence 11, Appli
6	1167	98.3	18596	24	US-09-631-275-140	Sequence 140, App
7	1167	98.3	18596	33	US-09-880-107-1590	Sequence 1590, App
8	1167	98.3	18596	36	US-09-954-531-124	Sequence 124, App
9	1167	98.3	18596	36	US-09-954-531-348	Sequence 348, App
10	1167	98.3	18596	36	US-09-967-768A-119	Sequence 119, App
11	1167	98.3	18597	25	US-09-658-659-8	Sequence 8, Appli
12	1167	98.3	18597	36	US-09-963-665-8	Sequence 8, Appli
13	1167	98.3	18597	36	US-09-963-677-8	Sequence 8, Appli
14	1167	98.3	18597	36	US-09-963-333-8	Sequence 8, Appli
15	335	28.2	1829	25	US-09-652-109-9676	Sequence 9676, Ap
16	335	28.2	1829	25	US-09-652-121-6686	Sequence 6686, Ap
17	335	28.2	1829	25	US-09-652-816-8870	Sequence 8870, Ap
18	335	28.2	1829	27	US-09-698-013-6995	Sequence 6995, Ap
19	335	28.2	1829	28	US-09-716-953-2244	Sequence 2244, Ap
20	335	28.2	1829	28	US-09-717-350-4489	Sequence 4489, Ap
21	335	28.2	1829	29	US-09-721-588-5305	Sequence 5305, Ap


```
QY 721 CCCTCGCGACGCTCTCTAGAGCGGGGCGCGCGACCCCGCGAGCAGAGAGCG 780
Db 797 CCCTGGCGACGCTCTCTAGAGCGGGGCGCGCGACCCCGCGAGCAGAGAGCG 856
QY 781 GAGCGGGAGCGCGCGGAAAAGGCGCGGAGGGGTCTGCGCCCGCGCGACTTGG 840
Db 857 GAGCGGGAGCGCGCGGAAAAGGCGCGGAGGGGTCTGCGCCCGCGCGACTTGG 916
QY 841 CTTGCTCTGCTCGCGCGGCGCACTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 900
Db 917 CTTGCTCTGCTCGCGCGGCGCACTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 976
QY 901 CTTGCTCTGCTCGCGCGGCGCACTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 960
Db 977 CTTGCTCTGCTCGCGCGGCGCACTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 1036
QY 961 TTGCGCGCGCGCGCAGAGCGGCGAGCGCGCGCTTGGCTTGGCTTGGCTTGG 1020
Db 1037 TTGCGCGCGCGCGCAGAGCGGCGAGCGCGCGCTTGGCTTGGCTTGGCTTGG 1096
QY 1021 TACTGCGGCGAGATCAACACATCTCTGCTGCGCGCTTGGCTTGGCTTGG 1080
Db 1097 TACTGCGGCGAGATCAACACATCTCTGCTGCGCGCTTGGCTTGGCTTGG 1156
QY 1081 ACCGCGACCCCTGTCGCTATTCGCGATGCGCGCGCTTGGCTTGGCTTGG 1140
Db 1157 ACCGCGACCCCTGTCGCTATTCGCGATGCGCGCGCTTGGCTTGGCTTGG 1216
QY 1141 GGCGCGCGCGCGCGGCGGAGAGGAGGCGCGCGCTTGGCTTGGCTTGG 1187
Db 1217 GGCGCGCGCGCGCGGCGGAGAGGAGGCGCGCGCTTGGCTTGGCTTGG 1263
```

RESULT 11

```
US-09-658-659-8
; Sequence 8, Application US/09658659
; GENERAL INFORMATION:
; APPLICANT: Stanton, Jr., Vincent P.
; TITLE OF INVENTION: GENE SEQUENCE VARIANCE IN GENES RELATED
; TO FOLATE METABOLISM HAVING UTILITY IN DETERMINING THE
; TITLE OF INVENTION: TREATMENT OF DISEASE
; FILE REFERENCE: 11926-015001
; CURRENT APPLICATION NUMBER: US/09/658,659
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 09/596,033
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 09/357,743
; PRIOR FILING DATE: 1998-07-20
; PRIOR APPLICATION NUMBER: 09/357,024
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: 60/093,484
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 18597
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 701, 13751
; OTHER INFORMATION: n = c or a
; NAME/KEY: misc_feature
; LOCATION: 716, 1293, 2401, 2429, 2618, 3083, 3125, 3635, 4256, 4898,
; LOCATION: 5062, 5167, 11069, 13298, 14479, 14730, 14796, 15344, 15450,
; LOCATION: 15503, 15590, 15840, 16149
; OTHER INFORMATION: n = a or g
; NAME/KEY: misc_feature
; LOCATION: 732, 1379, 1590, 2488, 3212, 5006, 11238, 11422, 11686,
; LOCATION: 12598, 13171, 13645, 13782, 13806, 13813, 14586, 14788,
; LOCATION: 15042, 15946, 15770
; OTHER INFORMATION: n = c or t
```

```
; NAME/KEY: misc_feature
; LOCATION: 1322, 1688
; OTHER INFORMATION: n = c or g
; NAME/KEY: misc_feature
; LOCATION: 2594, 11293, 16199, 16203
; OTHER INFORMATION: n = g or t
; NAME/KEY: misc_feature
; LOCATION: 3619
; OTHER INFORMATION: n = a or t
; NAME/KEY: misc_feature
; LOCATION: 14547
; OTHER INFORMATION: nucleotide in position 14547 is t, or absent
US-09-658-659-8
```

```
Query Match 98.3%; Score 1167; DB 25; Length 18597;
Best Local Similarity 99.5%; Pred. No. 4.5e-116;
Matches 1181; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 1 GATCGCGCGCTCACTCCAGCTCGGCTGAGAGCGAGACTCTGTCTCAAAAAAAA 60
Db 78 GATCGCGCGCTCACTCCAGCTCGGCTGAGAGCGAGACTCTGTCTCAAAAAAAA 137
QY 61 AAAAGGACCGCGAGCGCTCAAAACAAAAAAGCTCGAAAAAGCTTGTCTTTT 120
Db 138 AAAAGGACCGCGAGCGCTCAAAACAAAAAAGCTCGAAAAAGCTTGTCTTTT 197
QY 121 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 180
Db 198 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 257
QY 181 CGGATCTTGGCTCACTGCAACCTCTGCTCCAGCTTCAAGCAATCTTCTCTC 240
Db 258 CGGATCTTGGCTCACTGCAACCTCTGCTCCAGCTTCAAGCAATCTTCTCTC 317
QY 241 TCCCAAGTAGCCACCGCCAGCTAAATTTTGTANTTTTGTAGAGCGGGTTTCC 300
Db 318 TCCCAAGTAGCCACCGCCAGCTAAATTTTGTANTTTTGTAGAGCGGGTTTCC 377
QY 301 CATGTGTTCAGGCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
Db 378 CATGTGTTCAGGCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 437
QY 361 CAAAGTACTAGGATTAACAGCGCTGAGCCAGCGCTGAGCGCGCTGAGCGCG 420
Db 438 CAAAGTACTAGGATTAACAGCGCTGAGCCAGCGCTGAGCGCGCTGAGCGCG 497
QY 421 AGTAGAAGAGCTGCTATTAACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
Db 498 AGTAGAAGAGCTGCTATTAACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 556
QY 481 AAATGCAATCNCCTTATTTAGTTGTAGGAAAACAGATCTCAAAACAGAGTT 540
Db 557 AAATGCAATCNCCTTATTTAGTTGTAGGAAAACAGATCTCAAAACAGAGTT 616
QY 541 ACCGCGAGGAAAACGTTGGGAACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
Db 617 ACCGCGAGGAAAACGTTGGGAACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 676
QY 601 TTCCCAAGAGGCGCGAGTCTTCCGCGCAACCGCGCTGCTGCTGCTGCTGCTGCT 660
Db 677 TTCCCAAGAGGCGCGAGTCTTCCGCGCAACCGCGCTGCTGCTGCTGCTGCTGCT 736
QY 661 AAGACTCTCAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Db 737 AAGACTCTCAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 796
QY 721 CCCCTGCGCGCAGCTCTCTAGAGCGGGGCGCGCGCGCGCGCGCGCGCGCGCG 780
Db 797 CCCCTGCGCGCAGCTCTCTAGAGCGGGGCGCGCGCGCGCGCGCGCGCGCGCG 856
QY 781 GAGCGCGCGCGCGCGCGGAAAAGGCGCGCGGAGAGGGGCTCTTCCCGCGCGCT 840
Db 857 GAGCGCGCGCGCGCGCGGAAAAGGCGCGCGGAGAGGGGCTCTTCCCGCGCGCT 916
```


Qy 1021 TACCTGGGACATCCACATCCTCCGCTGGGCTCAGGAGGACGACCCGCGGC 1080
Db 1097 TACCTGGGACATCCACATCCTCCGCTGGGCTCAGGAGGACGACCCGCGGC 1156
Qy 1081 ACCGGCACCTCTCGGTATTCGGCATGACGCGCGCTACAGCTGAGAGTGACGCGCG 1140
Db 1157 ACCGGCACCTCTCGGTATTCGGCATGACGCGCGCTACAGCTGAGAGTGACGCGCG 1216
Qy 1141 GGGCCCTCGGGAACGGGTGGGGAAGGAGGAGGAGGCGCGCTGGGGA 1187
Db 1217 GGGCCCTCGGGAACGGGTGGGGAAGGAGGAGGAGGCGCGCTGGGGA 1263

RESULT 14

US-09-963-333-8

Sequence 8, Application US/09963333

GENERAL INFORMATION:

APPLICANT: Stanton, Jr., Vincent P.

TITLE OF INVENTION: THYMIDINE SYNTHASE GENE SEQUENCE VARIANCES

TITLE OF INVENTION: HAVING UTILITY IN DETERMINING THE TREATMENT

FILE REFERENCE: 11926-015002

CURRENT APPLICATION NUMBER: US/09/963,333

PRIOR FILING DATE: 2001-09-24

PRIOR FILING DATE: 2000-09-08

PRIOR FILING DATE: 2000-06-15

PRIOR FILING DATE: 1999-07-20

PRIOR FILING DATE: 1999-07-19

PRIOR FILING DATE: 1998-07-20

NUMBER OF SEQ ID NOS: 16

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 8

LENGTH: 18597

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc_feature

LOCATION: 701, 13751

OTHER INFORMATION: n = c or a

NAME/KEY: misc_feature

LOCATION: 716, 1293, 2401, 2429, 2618, 3083, 3125, 3635, 4256, 4898,

LOCATION: 5062, 5167, 11069, 13298, 14479, 14730, 14796, 15344, 15450,

LOCATION: 15503, 15590, 15840, 16149

OTHER INFORMATION: n = a or g

NAME/KEY: misc_feature

LOCATION: 732, 1379, 1590, 2488, 3212, 5006, 11238, 11422, 11686,

LOCATION: 12598, 13171, 13645, 13782, 13806, 13813, 14586, 14788,

LOCATION: 15042, 15546, 15770

OTHER INFORMATION: n = c or t

NAME/KEY: misc_feature

LOCATION: 1322, 1688

OTHER INFORMATION: n = c or g

NAME/KEY: misc_feature

LOCATION: 2594, 11293, 16199, 16203

OTHER INFORMATION: n = g or t

NAME/KEY: misc_feature

LOCATION: 3619

OTHER INFORMATION: n = a or t

NAME/KEY: misc_feature

LOCATION: 14547

OTHER INFORMATION: nucleotide in position 14547 is t, or absent

US-09-963-333-8

Query Match

Best Local Similarity 98.3%; Score 1167; DB 36; Length 18597;

Matches 1181; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

M nucleic - nucleic search, using sw model

run on: November 28, 2002, 21:45:20 ; Search time 70.1835 Seconds
(without alignments)
6513.474 Million cell updates/sec

file: US-09-963-333-7
effect score: 1187
sequence: 1 gatcgccgactgactcca.....gagggagcgcggtgggga 1187

scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

searched: 341543 seqs, 192557720 residues

total number of hits satisfying chosen parameters: 683086

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1167	98.3	18596	9	US-09-954-531-124
2	1167	98.3	18596	9	US-09-954-531-348
3	1167	98.3	18596	10	US-09-880-107-1590
4	1167	98.3	18596	10	US-09-967-768A-119
5	208.6	17.6	14012	9	US-09-819-994-3
6	206.8	17.4	9519	10	US-09-764-847-1056
7	206.8	17.4	9519	10	US-09-764-877-3534
8	202.8	17.1	5591	10	US-09-764-846-309
9	202.2	17.0	65464	9	US-09-859-888-3
10	198.4	16.7	53332	10	US-09-801-861-3
11	197.6	16.6	7017	10	US-09-764-877-3773
12	196.6	16.6	7537	10	US-09-764-869-1735
13	196.4	16.5	32248	10	US-09-764-864-1769
14	196.4	16.5	32248	10	US-09-764-877-3487
15	195.6	16.5	20522	10	US-09-764-877-3774
16	195	16.4	19616	10	US-09-764-877-3220
17	195	16.4	52216	10	US-09-747-810-1
18	194.8	16.4	32169	10	US-09-764-847-1963
19	194.6	16.4	32193	10	US-09-764-877-2623

Sequence 314, App
Sequence 3, Appli
Sequence 3, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 3, Appli
Sequence 3, Appli
Sequence 895, App
Sequence 244, App
Sequence 3525, Ap
Sequence 6, Appli
Sequence 202, App
Sequence 79, Appl
Sequence 2116, Ap
Sequence 1644, Ap
Sequence 1734, Ap
Sequence 201, App
Sequence 3250, Ap
Sequence 38, Appl
Sequence 5, Appli
Sequence 5, Appli
Sequence 281, App
Sequence 273, App
Sequence 81, Appl
Sequence 6, Appli

US-09-967-768A-314
US-09-759-359A-3
US-10-094-989-3
US-09-946-807-1
US-09-795-668-1
US-09-795-686-1
US-10-025-187-3
US-10-135-689-3
US-09-764-853-895
US-09-764-878-244
US-09-764-877-3525
US-09-835-232-6
US-09-764-878-202
US-09-969-708-79
US-09-954-456-2116
US-09-764-864-1644
US-09-764-869-1734
US-09-764-878-201
US-09-764-877-3250
US-09-927-602-38
US-09-739-457-5
US-09-982-091A-5
US-09-962-436-281
US-09-964-824A-273
US-09-898-779-81
US-09-967-736-6

ALIGNMENTS

RESULT 1

US-09-954-531-124

; Sequence 124, Application US/09954531

; Patent No. US20020165180A1

; GENERAL INFORMATION:

; APPLICANT: Weaver, Zoe

; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc

; FILE REFERENCE: 689290-77

; CURRENT APPLICATION NUMBER: US/09/954,531

; CURRENT FILING DATE: 2002-05-02

; PRIOR APPLICATION NUMBER: US/60/233,133

; PRIOR FILING DATE: 2000-09-18

; PRIOR APPLICATION NUMBER: US/60/234,009

; PRIOR FILING DATE: 2000-09-20

; PRIOR APPLICATION NUMBER: US/60/234,034

; PRIOR FILING DATE: 2000-09-20

; PRIOR APPLICATION NUMBER: US/60/234,509

; PRIOR FILING DATE: 2000-09-22

; PRIOR APPLICATION NUMBER: US/60/234,567

; PRIOR FILING DATE: 2000-09-22

; NUMBER OF SEQ ID NOS: 1392

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 124

; LENGTH: 18596

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-954-531-124

Query Match 98.3%; Score 1167; DB 9; Length 18596;
Best Local Similarity 99.2%; Pred.No. 1.2e-176;
Matches 1178; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

QY 1 GATCGCGCCACTGCTCAGCGCTGGGTGAGAGGAGAGAGCTGTCTCAAAAAA 60
Db 78 GATCGCGCCACTGCTCAGCGCTGGGTGAGAGGAGAGAGCTGTCTCAAAAAA 137
QY 61 AAAAAGACCGCGGCTCAAAACAAAAAACCCTCGAAAGCCCTGGCGGTCTTTT 120
Db 138 AAAAAGACCGCGGCTCAAAACAAAAAACCCTCGAAAGCCCTGGCGGTCTTTT 197
QY 121 TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT 180

592 AAAGAAAAGAAAGATGACGA-----GCAACGAAGTAAGATGAAGTTTTTTT 543
121 TTTTITTTTTTTTTTTTTTTTGGACAGCTCTCTGCGCCAGGCTGGAGTCAATGT 180
542 TTCTTTTTTTTTTTTTTTTGGAGGAGTCTCTCTGCGCCAGGCTGGAGTGGAGTGC 483
181 CGATCTTGCTCAGTCAAGCTCTGCTCCAGGTTCAAGCAATTTCTGCTCAGCC 240
482 ATGATCTTGCTCAGTCAAGCTCTGCTCCAGGTTCAAGCAATTTCTGCTCAGCC 423
241 TCCCAAGTA-----GCCACAGCCGCTCAATTTTGTANTTTTA 281
422 TCCCAAGTATGCTGCAATTAAGGCGCGCCACACCTCTGCTAATTTTATTTT 363
282 GTAGAGACGGGGTTTCCACATGTTTCCAGGCTGCTGCTGCTGCTGCTGCTGCTG 341
362 GTAGAGA-TGGGGTTTCCACATGTTTGGCCAGGCTGCTGCTGCTGCTGCTGCTG 304
342 TCCACCGGCTCGGCGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 400
303 TCGGCTGCTCAGCTCCCAAGTGTGGAATTAAGGCTGCTGCTGCTGCTGCTGCTG 245

Query Match 17.4%; Score 206.8; DB 10; Length 9519;
Best Local Similarity 82.8%; Pred. No. 5.7e-25;
Matches 255; Conservative 0; Mismatches 34; Indels 19; Gaps 1;

112 TTTTITTTTTTTTTTTTTTTTGGACAGCTCTCTGCGCCAGGCTGGAGTCAATGT 171
Db 2436 TATTAACTTTTTTTTTTTTTTTTGGACAGCTCTCTGCGCCAGGCTGGAG 2377

172 TACATGTCGGATCTTGGCTCACTGCAACCTCTGCTCCAGGTTCAAGCAATTTCT 231
Db 2376 TGCAGTAGCGTAATCTGGCTCACTGCAACCTCTGCTCCAGGTTCAAGCAATTTCT 2317

232 GCTCAGCTCCCAAGTAG-----CCACACGCGCCAGCTAATTTT 272
Db 2316 GCTCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2257

273 GTANTTTTGTAGAGACGGGGTTTCAACATGTTTCCAGGCTGGTCTGAACTCTCTG 332
Db 2256 GTATTTTGTAGAGACGGGGTTTCAACATGTTTCCAGGCTGGTCTGAACTCTCTG 2197

333 CTCAGTGATCAGCCGCTCGGCGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 392
Db 2196 CTCAGTGATCAGCCGCTCGGCGCCGCTGCTGCTGCTGCTGCTGCTGCTG 2137

393 CGTCCAGC 400
Db 2136 TGCCGAGC 2129

RESULT 8
US-09-764-846-309/c
; Sequence 309, Application US/09764846
; Patent No. US20020102638A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT212
; CURRENT APPLICATION NUMBER: US/09/764,846
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 309
; LENGTH: 5591
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-846-309

Query Match 17.1%; Score 202.8; DB 10; Length 5591;
Best Local Similarity 64.5%; Pred. No. 2.4e-24;
Matches 361; Conservative 0; Mismatches 168; Indels 31; Gaps 3;

1 GATCGCGCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60
Db 3620 GATCGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3561

592 AAAGAAAAGAAAGATGACGA-----GCAACGAAGTAAGATGAAGTTTTTTT 543
121 TTTTITTTTTTTTTTTTTTTTGGACAGCTCTCTGCGCCAGGCTGGAGTCAATGT 180
542 TTCTTTTTTTTTTTTTTTTGGAGGAGTCTCTCTGCGCCAGGCTGGAGTGGAGTGC 483
181 CGATCTTGCTCAGTCAAGCTCTGCTCCAGGTTCAAGCAATTTCTGCTCAGCC 240
482 ATGATCTTGCTCAGTCAAGCTCTGCTCCAGGTTCAAGCAATTTCTGCTCAGCC 423
241 TCCCAAGTA-----GCCACAGCCGCTCAATTTTGTANTTTTA 281
422 TCCCAAGTATGCTGCAATTAAGGCGCGCCACACCTCTGCTAATTTTATTTT 363
282 GTAGAGACGGGGTTTCCACATGTTTCCAGGCTGCTGCTGCTGCTGCTGCTGCTG 341
362 GTAGAGA-TGGGGTTTCCACATGTTTGGCCAGGCTGCTGCTGCTGCTGCTGCTG 304
342 TCCACCGGCTCGGCGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 400
303 TCGGCTGCTCAGCTCCCAAGTGTGGAATTAAGGCTGCTGCTGCTGCTGCTGCTG 245

Query Match 17.4%; Score 206.8; DB 10; Length 9519;
Best Local Similarity 82.8%; Pred. No. 5.7e-25;
Matches 255; Conservative 0; Mismatches 34; Indels 19; Gaps 1;

112 TTTTITTTTTTTTTTTTTTTTGGACAGCTCTCTGCGCCAGGCTGGAGTCAATGT 171
Db 2436 TATTAACTTTTTTTTTTTTTTTTGGACAGCTCTCTGCGCCAGGCTGGAG 2377

172 TACATGTCGGATCTTGGCTCACTGCAACCTCTGCTCCAGGTTCAAGCAATTTCT 231
Db 2376 TGCAGTAGCGTAATCTGGCTCACTGCAACCTCTGCTCCAGGTTCAAGCAATTTCT 2317

232 GCTCAGCTCCCAAGTAG-----CCACACGCGCCAGCTAATTTT 272
Db 2316 GCTCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2257

273 GTANTTTTGTAGAGACGGGGTTTCAACATGTTTCCAGGCTGGTCTGAACTCTCTG 332
Db 2256 GTATTTTGTAGAGACGGGGTTTCAACATGTTTCCAGGCTGGTCTGAACTCTCTG 2197

333 CTCAGTGATCAGCCGCTCGGCGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 392
Db 2196 CTCAGTGATCAGCCGCTCGGCGCCGCTGCTGCTGCTGCTGCTGCTGCTG 2137

393 CGTCCAGC 400
Db 2136 TGCCGAGC 2129

RESULT 7
US-09-764-877-3534/c

QY 61 AAAAGAGCCGACGAGGCTCAACAAAAAACCCTCGAAAAAGCCCTGGCGGCTCTTTTTTTT 120
Db 3560 AAAAAGAGCCAGTTCCTCCAAAAGATCAACCGTTACTTATCAGGAACA----- 3510
QY 121 TTTTTTTTTTTTTTTTTTGGGACAGCTTCTGCTGTGCGCCAGAGCTGGATCAATGGT 180
Db 3509 -TTTTTTTTTTTTTTTTTGGACGAGCTATCACTGTCTATCCAGGCTGGAGACAATGGC 3452
QY 181 CGGATCTGGCTCACTGAACCTCTGCTCCAGGTTCAAGCAATTTCTGCGCTCAGCC 240
Db 3451 GCGATCTGGCTTGTGCAACCTCGGCTCTGCGGTTCAAGCAATTTCTGCGCTCAGTC 3392
QY 241 TCCCAAGTA-----GCCACACGCCAGCTAATTTTTGTATTTTTA 281
Db 3391 TCCTGAGTAGCTGGGAAATACAGACATGTGCGCACCAACACAGCTCATTTTGTGATTTTA 3332
QY 282 GTAGAGAGGGGGTTTCAACATGTGTCCAGGCTGTCTNGAACTCTGACCTCAGGTGA 341
Db 3331 GTATAGAC-GGGGTTTGGCAATGTGGCCAGGCTGTCTCAAACTCCTGACCTCAGGTGA 3273
QY 342 TCACCCCGCTCGGCGCCGCCAAAGTACTAGGATTTACAGGCTGAGCCACCGGCTCAGCG 401
Db 3272 TCTACCGCTCTGGCTCTCCAAAGTGTAGCATTTACAGGCTGAGCCAAATGCGCGCC 3213
QY 402 CCCTGGCGGTTTTTAATCAAGTAGAAAAGCTGCAATATACCATTGCTTGTGCTGCTTC 461
Db 3212 TCATATCCAGGAACATTTTATCAAACTCCAAATTTCAAGTATAGTATTTAAAGTA 3153
QY 462 AGTGAGAACGAAGATGGAATGCAATCATTATTAGTTGTAGGAACAGATCTCAA 521
Db 3152 TGAATCAAGAGACTAGAAAAAGGATATCCATTCAGCATTTAGACAAAGTTATACCTCC 3093
QY 522 CAGCATTTTGTNGACAGA 541
Db 3092 ATTTTGTCTGGGTGAAGGA 3073

RESULT 9
US-09-859-888-3/c
; Sequence 3, Application US/09859888
; Patent No. US20020173459A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL001239
; CURRENT APPLICATION NUMBER: US/09/859,888
; CURRENT FILING DATE: 2001-05-18
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 65464
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(65464)
; OTHER INFORMATION: n = A,T,C or G
US-09-859-888-3

Query Match 17.0%; Score 202.2; DB 9; Length 65464;
Best Local Similarity 80.5%; Pred. No. 3.3e-24;
Matches 268; Conservative 0; Mismatches 45; Indels 20; Gaps 2;

QY 106 GCGGTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGGGACAGTCTGTCTGTGCGCCAGG 165
Db 51995 GGGGTACTTATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGGAGGGAGTCTTGTCTGTGCGCCAGG 51936
QY 166 CTGGAGTCAATGGTGGGATCTGGCTCACTGCACTCTCCCTCCAGGTTCAAGCAAT 225
Db 51935 CTGGAGTCAATGGGAGTCTGGGCTCACTGCACTCTCCCTCCAGGTTCAAGCAAT 51876

QY 226 TCTTCTGCTCAGCTCCCAAGTAG-----CCACACGCCAGCTA 266
Db 51875 TCTCTGCTCAGCTCCCAAGTAGCTGGGATTTACAGGATGCAACACGCCCTGGCTA 51816
QY 267 ATTTTGTATTTTATAGTAGAGACGGGGTTTTCACCATGTTGTCCAGGCTGTCTNGAACT 326
Db 51815 ATTTTGTATTTTATAGTAGAGAC-AGGGTTTGGCAATGTGGCCAGGCTGTCTNGAACT 51757
QY 327 CTTGACCTCAGGTGATCCACCGGCTCGGCCCCCAAAAGTACTAGGATTTACAGGCTGAG 386
Db 51756 CTTGACCTCAATGATCCACCGGCTCGGCCCTCCAAAATGTGGGATTTACAGGCTGAG 51697
QY 387 CCACGCGTCCAGGCGCTCGGCGGTTTTTAATC 419
Db 51696 CCACGCGCGCGGCGCTCGGCGGTTTTTAATC 51664
RESULT 10
US-09-801-861-3/c
; Sequence 3, Application US/09801861
; Patent No. US20020119544A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001098
; CURRENT APPLICATION NUMBER: US/09/801,861
; CURRENT FILING DATE: 2001-03-09
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 53332
; TYPE: DNA
; ORGANISM: Human
US-09-801-861-3

Query Match 16.7%; Score 198.4; DB 10; Length 53332;
Best Local Similarity 78.2%; Pred. No. 1.3e-23;
Matches 258; Conservative 0; Mismatches 53; Indels 19; Gaps 1;

QY 110 TCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGGGACAGTCTGTCTGTGCGCCAGGCTGG 169
Db 7657 TCTTTTTTGTTTTTTTTTTTTTTTTTTTTTTTTTCAGACAGATCTGCTCTGTGCGCCAGGCTGG 7598
QY 170 AGTAAATGTCGGATCTTGGCTCACTGCAACCTCTGCTCCAGGTTCAAGCAATTTCTT 229
Db 7597 AGTACAGTGGACAGATCTTGGCTCACTGCAACCTCCACCTCTGSGTTCAAGCAATTTCTC 7538
QY 230 CTGCTCTCAGCTCCCAAGTAG-----CCACACGCCAGCTAATTT 270
Db 7537 CTGCTCAGCTCCCAAGAGCTGGGATTTACAGGGGACCCCAACACCTGCTAATTT 7478
QY 271 TTGTATTTTATAGTAGAGACGGGGTTTTCACCATGTTGTCCAGGCTGTCTNGAACTCTG 330
Db 7477 TTGTATTTTATAGTAGAGATGGGGTTTTCACCATTTTGGCCAGGCTGTCTGCAATTTCT 7418
QY 331 ACCTCAGGTGATCCACCGGCTCGGCCCCCAAAAGTACTAGGATTTACAGGCTGAGCCAC 390
Db 7417 ACCTCAGGTGATCTGCCACGTCAGGCTCCCAAGTGTGGGATTTACAGGCTGAGCCAC 7358
QY 391 CGGCTCCAGCGCTCGGCGGTTTTTAATCA 420
Db 7357 CAAGCCAGCTTGTAGTGTCTTTTAACA 7328

RESULT 11
US-09-764-877-3773
; Sequence 3773, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

Db	438	CAAAAGTACTAGAGATTACAGCGGTGAGCCACCGCGTCCAGGCGCCTTGCGGGTTTAAATCA	4397
QY	421	AGTAGAAAAGCTGCATTATACCACTTGTCTTCNGTTCGNTTCAGTGAAGACAGAAATCG	480
Db	498	AGTAGAAAAGCTGCATTATACCACTTGTCTCGGTTC-TTCAGTGAGAACGAGAAATGG	556
QY	481	AAATGCAAAATGNCCTTTATTAGTTTATAGAAAACAGATCTCAAAACAGACAGTTTGTGACAAAG	540
Db	557	AAATGCAAAATCCCTTTATTAGTTTATAGAAAACAGATCTCAAAACAGCAGTTTTGTGACAAAG	616
QY	541	ACCGCTGGAAAACGTGGGAACTGTGCTGCTGTATAGAGAGGCGCGGTACGACAGACGG	600
Db	617	ACCGCAGGAAAACGTGGGAACTGTGCTGCTGTATAGAGAGGCGCGGTACGACAGACGG	676
QY	601	TTCCGAAAAGGGCGAGTCCCTTCCGNGCCACCGACCTGCTGCTGACAGTTCCCGGGTTTCCT	660
Db	677	TTCCGAAAAGGGCGAGTCCCTTCCGACCCACCGACCTGATCCAGGTTCCCGGGTTTCCT	736
QY	661	AAGACTCTCAGCTGTGAGCCCTCGGGCTTCGTTCTGTGACCAACCGCTGGCTCTCGGTTTC	720
Db	737	AAGACTCTCAGCTGTGAGCCCTCGGGCTTCGTTCTGTGACCAACCGCTGGCTCTCGGTTTC	796
QY	721	CCCTCTGCGCAAGCTCTCTAGAGCGGGGGCCGCGCGAACCCCGCGCAGGAAAGAGGCG	780
Db	797	CCCTCTGCGCAAGCTCTCTAGAGCGGGGGCCGCGCGAACCCCGCGCAGGAAAGAGGCG	856
QY	781	GAGCGCGGACCGCGCGGGAAAAGCGCGCGGAAAGGGGTCTGTGCACCGCGCCACTTGG	840
Db	857	GAGCGCGGACCGCGCGGGAAAAGCGCGCGGAAAGGGGTCTGTGCACCGCGCCACTTGG	916
QY	841	CCTGCTCCGTCGCCGCGCGCACCTTGGCGTTCCTCGCTCCGCGCGCGCACCTTGCCTTG	900
Db	917	CCTGCTCCGTCGCCGCGCGCACCTTGGCGTTCCTCGCTCCGCGCGCGCACCTTGCCTTG	976
QY	901	CCTCGTCCCGCGCGCGCGCCATGCTGTGCGCGGTCTGGAGCTGCGCGCGCGCGCC	960
Db	977	CCTCGTCCCGCGCGCGCGCCATGCTGTGCGCGGTCTGGAGCTGCGCGCGCGCGCC	1036
QY	961	TTGCGCCCGCGCGCACAGGAGCGGAGCGCCGAGCGCGCTCCCGCGCGGGAGTGTGAG	1020
Db	1037	TTGCGCCCGCGCGCACAGGAGCGGAGCGCCGAGCGCGCTCCCGCGCGGGAGTGTGAG	1096
QY	1021	TACCTGSGCGAGATCCAAACATCTTCGCTGCGGGCTCAGGAAAGCACACCGCACGGCC	1080
Db	1097	TACCTGSGCGAGATCCAAACATCTTCGCTGCGGGCTCAGGAAAGCACACCGCACGGCC	1156
QY	1081	ACCGGACCCCTCTCGGTATTTCGGCATGACGGCGGCTTACAGCCTTGAGAGTGAOCGCCG	1140
Db	1157	ACCGGACCCCTCTCGGTATTTCGGCATGACGGCGGCTTACAGCCTTGAGAGTGAOCGCCG	1216
QY	1141	GGCCCTCTGGGAGCGGTTGGCGGAAAGAGGAGGCGCGCTTGGGGA	1187
Db	1217	GGCCCTCTGGGAGCGGTTGGCGGAAAGAGGAGGCGCGCTTGGGGA	1263

RESULT 2

US-09-089-195-1
; Sequence 1, Application US/09089195
; Patent No 6087489

FACEBOOK NO. 0067483
GENERAL INFORMATION:

```

/ GENERAL INFORMATION:
/ APPLICANT: Dean, Nicholas M.
/
/ TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDE MODULATION
/
/ TITLE OF INVENTION: OF HUMAN THYMIDYLATE SYNTHASE EXPRESSION
/
/ NUMBER OF SEQUENCES: 28
/
/ CORRESPONDENCE ADDRESS:

```

```

CORRESPONDENCE ADDRESS:
ADDRESSER: Law Offices of Jane Massey Licata
STREET: 66 East Main Street
CITY: Marlton
STATE: NJ
COUNTRY: USA
ZIP: 08053
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

```

RESULT 3

US-09-367-007C-38
; Sequence 38, Application US/09367007C

; FACILE NO. 841858)
: GENERAL INFORMATION:

; APPLICANT: Bertino, Joseph R.
 ; APPLICANT: Banerjee, Debabrata
 ; APPLICANT: Tong, Youzhi

Query Match	25.2%;	Score 299.4;	DB 3;	Length 1536;
Best Local Similarity	99.7%;	Pred. No. 2.7e-47;		
Matches 300; Conservative	0;	Mismatches 1;	Indels 0;	

QY	833	CCAATTGGGCTGCTCTCCGTTCCCGCGCGGCACTTGGGCTGCTCTCGTCCGCGCGCGCAC	892
Db	14	CCAATTGGGCTGCTCTCCGTTCCCGCGCGGCACTTGGGCTGCTCTCGTCCGCGCGCGCAC	73
QY	893	TTGGCTTGCTTCCGTTCCCGCGCGCGCGGCAATGCTGTGGCCGGCTCGGAGCTGCCGC	952
Db	74	TTGGCTTGCTTCCGTTCCCGCGCGCGCGGCAATGCTGTGGCCGGCTCGGAGCTGCCGC	133
QY	953	GCGGGCCCTTGGCCCCCGCGCGCACAGGAGCGGAGCGCCGAGCGCGCTCCGCGCACGGGG	1012
Db	134	GCGGGCCCTTGGCCCCCGCGCGCACAGGAGCGGAGCGCCGAGCGCGCTCCGCGCACGGGG	193
QY	1013	AGCTGACGTACTGGGGCAGATCCAAACATCTTCGCTGGCGGGCTCAGGAAGGACGACC	1072
Db	194	AGCTGACGTACTGGGGCAGATCCAAACATCTTCGCTGGCGGGCTCAGGAAGGACGACC	253
QY	1073	GCACGGGCACCGGCACCCCTGTGGGTATTTCGGCATTCAGGCGCGGCTACAGCTTGAGAGTG	1132
Db	254	GCACGGGCACCGGCACCCCTGTGGGTATTTCGGCATTCAGGCGCGGCTACAGCTTGAGAGTG	313
QY	1133	A 1133	
Db	314	A 314	

```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE:
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/089,195
FILING DATE: herewith
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Jane Massey Licata
REGISTRATION NUMBER: 32,257
REFERENCE/DOCKET NUMBER: ISPH-0309
TELECOMMUNICATION INFORMATION:
TELEPHONE: (609) 810-1515
TELEFAX: (609) 810-1454
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1536 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Unknown
PUBLICATION INFORMATION:
AUTHORS: Takeishi, K.
AUTHORS: Kaneda, S.
AUTHORS: Ayusawa, D.
AUTHORS: Shimizu, K.
AUTHORS: Gotoh, O.
AUTHORS: Seno, T.
TITLE: Nucleotide sequence of a functional cDNA
TITLE: for human thymidylate synthase
JOURNAL: Nucleic Acids Res.
VOLUME: 13
ISSUE: 6
PAGES: 2035-2043
DATE: 25-MAR-1985
US-09-089-195-1

```

APPLICANT: Liu-Chen, Xinyue
TITLE OF INVENTION: Mutants of Thymidylate Synthase and Uses Thereof
FILE REFERENCE: D5978
CURRENT APPLICATION NUMBER: US/09/367,007C
CURRENT FILING DATE: 1999-10-15
PRIOR APPLICATION NUMBER: PCT/US98/02145
PRIOR FILING DATE: 1998-01-03
NUMBER OF SEQ ID NOS: 39
SEQ ID NO 38
LENGTH: 1536
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Wild type human thymidylate synthase cDNA
OTHER INFORMATION: (Genbank Accession number IM 001071)
JS-09-367-007C-38

Query Match 25.2%; Score 299.4; DB 4; Length 1536;
Best Local Similarity 99.7%; Pred. No. 2.7e-47;
Matches 300; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
2Y 833 CAATTGGCTGCTCTCGTCCGCGCGGACCTGTCCTCGTCCGCGCGGAC 892
Db 14 CCATTGGCTGCTCTCGTCCGCGCGGACCTGTCCTCGTCCGCGCGGAC 73
2Y 893 TTCCCTGCTGCTCTCGTCCGCGCGGACCTGTCCTCGTCCGCGCGGAC 952
Db 74 TTCCCTGCTGCTCTCGTCCGCGCGGACCTGTCCTCGTCCGCGCGGAC 133
QY 953 GCCGCGCTTTCGCGCGGACCTGTCCTCGTCCGCGCGGACCTGTCCT 1012
Db 134 GCCGCGCTTTCGCGCGGACCTGTCCTCGTCCGCGCGGACCTGTCCT 193
QY 1013 AGCTGAGTACCTGCGGCGGACCTGTCCTCGTCCGCGCGGACCTGTC 1072
Db 194 AGCTGAGTACCTGCGGCGGACCTGTCCTCGTCCGCGCGGACCTGTC 253
QY 1073 GCAGGCGACCGGACCTGTCCTCGTCCGCGCGGACCTGTCCTCGTCC 1132
Db 254 GCAGGCGACCGGACCTGTCCTCGTCCGCGCGGACCTGTCCTCGTCC 313
QY 1133 A 1133
Db 314 A 314
RESULT 4
US-09-347-878-29
; Sequence 29, Application US/09347878C
; Patent No. 6376210
; GENERAL INFORMATION:
; APPLICANT: Yuan, Chong
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR ASSAYING ANALYTES
; FILE REFERENCE: 25885-1651
; CURRENT APPLICATION NUMBER: US/09/347,878C
; CURRENT FILING DATE: 1999-07-06
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 29
; LENGTH: 942
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Human thymidylate synthase gene: exons 1-8
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: D00596/GenBank
US-09-347-878-29
Query Match 17.5%; Score 207.4; DB 4; Length 942;
Best Local Similarity 99.5%; Pred. No. 2.6e-30;
Matches 208; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 925 ATGCTGTGGCGGCTCGGAGCTGCGCGCGGCGCTTGGCCCGCGCGGAGGCGG 984

Db 1 ATCCCTGTGGCGGCTCGGAGCTGCGCGCGGCGCTTGGCCCGCGCGAGGCGG 60
QY 985 GAGCGCGAGCGCGCTCCGCGCGGACCTGTCCTCGTCCGCGCGGACCTG 1044
Db 61 GAGCGCGAGCGCGCTCCGCGCGGACCTGTCCTCGTCCGCGCGGACCTG 120
QY 1045 CTCGCTGCGGCGTTCAGGAAGGACGACCGCGCGGACCTGTCCTCGT 1104
Db 121 CTCGCTGCGGCGTTCAGGAAGGACGACCGCGCGGACCTGTCCTCGT 180
QY 1105 ATGCGGCGGCTACAGCTGAGAGTGA 1133
Db 181 ATGCGGCGGCTACAGCTGAGAGATGA 209

RESULT 5

US-09-780-049-18
; Sequence 18, Application US/09780049
; Patent No. 6465250
; GENERAL INFORMATION:
; APPLICANT: Brett P. Monia
; APPLICANT: Jacqueline Wyatt
; TITLE OF INVENTION: ANTISENSE MODULATION OF PROTEIN PHOSPHATASE 2 CATALYTIC SUBUNIT 1
; TITLE OF INVENTION: EXPRESSION
; FILE REFERENCE: KTS-0134
; CURRENT APPLICATION NUMBER: US/09/780,049
; CURRENT FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 96
; SEQ ID NO 18
; LENGTH: 40000
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-780-049-18

Query Match 16.7%; Score 198.6; DB 4; Length 40000;
Best Local Similarity 79.1%; Pred. No. 1.8e-28;
Matches 268; Conservative 0; Mismatches 51; Indels 20; Gaps 2;
QY 104 TGGCGGCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 163
Db 7055 TTGTGGGCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 7114
QY 164 GGCTGGAGTACAATGGTTCGGATCTTGGCTCACTGCAACCTCTGCTCCAGGTTCAAGCA 223
Db 7115 GGCTGGAGTCAATGGTTCGGATCTTGGCTCACTGCAACCTCTGCTCCAGGTTCAAGCA 7174
QY 224 ATCTTCTGCTCAGCTCCCAAGTA-----GCCACCAAGCCCGCAGC 264
Db 7175 GTTCTCTGCTCAGCTCCCAAGTA-----GCCACCAAGCCCGCAGC 7234
QY 265 TAAATTTTGTANTTTTAGTAGAGAGCGGGTTTCCACCATGTTGTCAGGCTGCTGAA 324
Db 7235 TAAATTTTGTANTTTTAGTAGAGAGC-AGGGTTTCCACCATGTTGTCAGGCTGCTGAA 7293
QY 325 CTCCTGACCTCAGGTGATCCACCCCTCGCGCGCGCGCGCGCGCGCGCG 384
Db 7294 CTCCTGACCTCAGGTGATCCACCCCTCGCGCGCGCGCGCGCGCGCGCG 7353
QY 385 AGCCACCGGCTCAGCGCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCG 423
Db 7354 AGCCACCGGCTCAGCGCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCG 7392

RESULT 6

US-09-146-053-6/c
; Sequence 6, Application US/09146053A
; Patent No. 6399349
; GENERAL INFORMATION:
; APPLICANT: Ryan, James W.
; APPLICANT: Sprinkle, Terry Joe Curtis
; APPLICANT: Venema, Richard C.

;; TITLE OF INVENTION: Human Aminopeptidase P Gene
;; FILE REFERENCE: MCG103
;; CURRENT APPLICATION NUMBER: US/09/146,053A
;; CURRENT FILING DATE: 1998-09-02
;; EARLIER APPLICATION NUMBER: 60/057,854
;; EARLIER FILING DATE: 1997-09-02
;; NUMBER OF SEQ ID NOS: 7
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 6
;; LENGTH: 45546
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-146-053-6

Query Match 16.6%; Score 197.4; DB 4; Length 45546;
Best Local Similarity 82.3%; Pred. No. 3e-28; 33; Indels 20; Gaps 1;
Matches 246; Conservative 0; Mismatches 33; Indels 20; Gaps 1;
QY 126 TTTTCTTTTGGGACAGTCTTGTCTGTGCGCCAGGCTGGAGTACATGTCGGAT 185
DB 1740 TTTTCTTTTGGGACAGTCTTGTCTGTGCGCCAGGCTGGAGTACATGTCGGAT 1681
QY 186 CTTGGCTCACTGCACTCTGCTCCAGGTTCAAGCAATTTCTTGTCTGAGCTCCCA 245
DB 1680 CTCAGCTCACTGCACTCTGCTCCAGGTTCAAGCAATTTCTTGTCTGAGCTCCCA 1621
QY 246 AGTAG-----CCACCACGCCAGCTAAATTTTGTANTTTTAGTAG 285
DB 1620 AGTAGCTGGAGTACAGGATGACCCACACAGCTAAATTTTGTANTTTTAGTAG 1561
QY 286 AGACGGGGTTTCAACATTTGTCCAGGCTGGTCTGAACTCTCTGAGCTCAGGTATCCA 345
DB 1560 AGATGGGGTTTCACTATATTGGCCAGGCTGGTCTGAACTCTCTGAGCTCAGGTATCCA 1501
QY 346 CCGCTCGGCCCCCAAGTACTAGATTTACAGGCTGAGCAGCGGTCCAGCGCC 404
DB 1500 CCGGCCCGGCTCCCAAGTGTGGGATCACAGGCTGAGCAGCGGTCCAGCGCC 1442

RESULT 7
US-09-754-250-3
; Sequence 3, Application US/09754250
; Patent No. 6376225
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al
; TITLE OF INVENTION: ISOLATED HUMAN PHOSPHODIESTERASE
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: PHOSPHODIESTERASE PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL001063
; CURRENT APPLICATION NUMBER: US/09/754,250
; CURRENT FILING DATE: 2001-01-05
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 111282
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(111282)
; OTHER INFORMATION: n = A,T,C or G
US-09-754-250-3

Query Match 16.3%; Score 193.2; DB 4; Length 111282;
Best Local Similarity 81.0%; Pred. No. 2e-27; 40; Indels 20; Gaps 2;
Matches 256; Conservative 0; Mismatches 40; Indels 20; Gaps 2;
QY 104 TGGCGGCTTTTGGGACAGTCTTGTCTGAGCTCTTGTCTGTCGGCCA 163
DB 18301 TGGTCTTCTTTCTTCTTTTGGGACAGTCTTGTCTGTCGGCCA 18360
QY 164 GCCTGAGTACAATGGTCGGATCTTGGCTCACTGCAACCTCTGCTCCAGGTTCAAGCA 223

DB 18361 GGCTGGAGTGCAGTGGTGGATCTTGTCTCACTGCAACCTCCACCTCTCGGTTCAAGTG 18420
QY 224 ATTCTTCTGCTCAGCTCCCAAGTA-----GCCACACGCCCGCAGC 264
DB 18421 ATTCTTCTGCTCAGCTCCCAAGTAGCTGGATCACAGGCACCTGCCAGCATGCCGGC 18480
QY 265 TAATTTTGTANTTTTAGTAGAGAGCGGGTTTCAACATTTGTCCAGGCTGGTCTNGAA 324
DB 18481 TAATTTTGTACTTTTGGTGTAGAGAC-GGGGTTTACCATATTGGCCAGGCTGGTCTCGAA 18539
QY 325 CTCCTGACCTCAGGTGATCCACCGCTCGGCCCCCAAGTACTAGGATTACAGGCTG 384
DB 18540 CTCCTGCCCTCAGGTGATCTGCCGCTTACCTCCCAAGTGTGGGATTACAGGCTG 18599
QY 385 AGCCACCGCTCCAGC 400
DB 18600 AGCCACCGCTCCAGC 18615
RESULT 8
US-09-729-995-3/c
; Sequence 3, Application US/09729995
; Patent No. 6426206
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000904
; CURRENT APPLICATION NUMBER: US/09/729,995
; CURRENT FILING DATE: 2000-12-06
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 29629
; TYPE: DNA
; ORGANISM: Human
US-09-729-995-3

Query Match 16.2%; Score 192; DB 4; Length 29629;
Best Local Similarity 81.6%; Pred. No. 2.8e-27;
Matches 253; Conservative 0; Mismatches 37; Indels 20; Gaps 2;
QY 110 TCTTTTCTTTTGGGACAGTCTTGTCTGTCGCCAGGCTGG 169
DB 11939 TTTTCTTTTGGTCTTTTGGACAGATCTCTGTCTCCCGAGGCTGG 11880
QY 170 AGTCAATGTGGATCTTGGCTCACTGCAACCTCTGCTCCAGGTTCAAGCAATCTT 229
DB 11879 AGTCAATGTGGATCTGCGCTCACTGCAACCTCTCCAGGTTCAAGCAATCTC 11820
QY 230 CTGCTCAGCTCCCAAGTA-----GCCACACGCCCGCAGCTAATTT 270
DB 11819 CTGCTCAGCTCCCGAGTAGTGGGATACAGGTGTCTGCCACCGCCGCTACTTT 11760
QY 271 TTGTANTTTTAGTAGAGAGCGGGTTTCAACATTTGTCCAGGCTGGTCTNGAATCTCG 330
DB 11759 TTGTANTTTTAGTAGAGAC-AGGTTTTCCTATTTGGTCAGGCTGGTCTTGAATCTCG 11701
QY 331 ACCTCAGGTGATCCACCGCTCGGCCCCCAAGTACTAGGATTACAGGCTGAGCCAC 390
DB 11700 ACCTCAGGTGATCCACCGCTCGGCCCTCCCAAGTCTGGGATTACAGGCTGAGCCAC 11641
QY 391 CGCGTCCAGC 400
DB 11640 CGCGTCCAGC 11631

RESULT 9
US-09-608-285A-42
; Sequence 42, Application US/09608285A
; Patent No. 6335013
; GENERAL INFORMATION:

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

M nucleic - nucleic search, using sw model

run on: November 28, 2002, 17:24:19 ; Search time 234.959 Seconds
(without alignments)
11376.979 Million cell updates/sec

Title: US-09-963-333-7
Perfect score: 1187
Sequence: 1 gatcgccactgcactcca.....ggggggcgcggtg999ga 1187

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq_101002.*
1: /SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1167	98.3	18596	22	AAF31109
2	1167	98.3	18596	22	AAAC91215
3	1167	98.3	18596	24	ABN95092
4	1167	98.3	18596	24	ABK43334
5	1167	98.3	18596	24	ABL62854
6	1167	98.3	18596	24	ABL63078
7	1167	98.3	18596	24	ABL67927
8	770.6	64.9	1186	24	ABK40090
9	710	59.8	1186	24	ABK40089

10	299.4	25.2	1536	24	ABK43335	Human Thymidylate
11	288.8	24.3	3298	24	AAS94945	Human DNA sequence
12	226.6	19.1	764	24	ABQ39526	Oligonucleotide fo
13	226.6	19.1	764	24	ABQ39527	Oligonucleotide fo
14	223.2	18.8	764	24	ABQ39528	Oligonucleotide fo
15	223.2	18.8	764	24	ABQ39529	Oligonucleotide fo
16	216.2	18.2	136284	24	ABK83575	Human cDNA differe
17	207.4	17.5	1539	23	AAS94960	DNA encoding novel
18	206.8	17.4	9519	22	AAI37169	Human musculoskele
19	206.8	17.4	9519	22	AAI37169	Human musculoskele
20	206.8	17.4	56632	22	AAK65581	Human immune/haema
21	204.8	17.3	9731	22	ABL04943	Human reproductive
22	204.8	17.3	9731	22	ABL04943	Human testicular a
23	204.6	17.2	11453	22	AAK73272	Human immune/haema
24	204.6	17.2	11453	22	AAK73272	Human immune/haema
25	204.2	17.2	37449	22	AAK66874	Human immune/haema
26	204.2	17.2	56743	22	AAK68202	Human immune/haema
27	204.2	17.2	56743	22	AAK81760	Human immune/haema
28	203.6	17.2	20601	22	AAK79760	Human nervous syst
29	202.8	17.1	5591	22	ABAI5206	Genomic sequence #
30	202.8	17.1	5591	22	ABAI5206	Genomic sequence #
31	202.6	17.1	22916	22	AAK65305	Human immune/haema
32	201.8	17.0	10820	22	ABAI6598	Human nervous syst
33	201.8	17.0	122888	24	ABK83569	Human cDNA differe
34	200	16.8	9060	24	AAS19272	Human genomic DNA
35	200	16.8	9060	24	AAS19272	Human generic DNA
36	199.4	16.8	32169	22	ABAI4358	Human nervous syst
37	198.6	16.7	32173	22	ABA07767	Human ovarian and
38	198.6	16.7	32173	22	ABA07767	Human reproductive
39	198.6	16.7	38928	22	AAK68452	Human immune/haema
40	198.6	16.7	109906	24	ABK94411	DNA encoding endot
41	197.8	16.7	846	22	AAS44955	CDNA encoding nove
42	197.8	16.7	20050	24	ABL53875	Human immunoglobul
43	197.6	16.6	7017	22	AAI37408	Human musculoskele
44	197.6	16.6	50000	21	AAA96365	Polymorphic repeat
45	197.4	16.6	45546	20	AAK23520	Human kidney amino

ALIGNMENTS

RESULT 1
AAF31109
ID AAF31109 standard; cDNA; 18596 BP.
XX
AC AAF31109;
XX
DT 27-APR-2001 (first entry)
XX
Thymidylate synthase coding sequence.
XX
Analyte-binding enzyme; analyte analysis; ss.
XX
Homo sapiens.
XX
WO200102600-A2.
XX
11-JAN-2001.
XX
30-JUN-2000; 2000WO-US18057.
XX
06-JUL-1999; 99US-0347878.
XX
06-DEC-1999; 99US-0457205.
XX
(GEAT) GEN ATOMICS.
XX
Yuan C;
XX
WPI; 2001-071583/08.
XX
Assaying method, useful for prognosis and diagnosis of disease,
XX
comprises contacting sample with a mutant analyte-binding enzyme and
XX
detecting binding -

X	Q	Sequence 18596 BP; 4521 A; 3991 C; 4479 G; 5605 T; 0 other;
		Query Match 99.3%; Score 1167; DB 22; Length 18596;
		Best Local Similarity 99.2%; Pred. No. 1.1e-193;
		Matches 1178; Conservative 0; Mismatches 8; Indels 1; Gaps
y		1 GATCGCGCACCTGCATCTCCAGCTGGGTGAGAGACGAGACTCTGTCCTCAAAAAA 60
b		78 GATCGCGCCACTGCATCTCCAGCTGGGTGAGAGACGAGACTCTGTCCTAAAAA 137
y		61 AAAAGACCGCCAGGGCTCAAAACAAAAACCTCGAAAAACCCCTGCGCGTCTTTT 120
b		138 AAAAGACCGCCAGGGCTCAAAACAAAAACCTCGAAAAACCCCTGCGCGTCTTT 197
y		121 TTTTTTTTTTTTTTTTTTTTTTGGGACAGTCTTGTCTGTGCCACAGGCTGAGTGA 180
b		198 TTTTTTTTTTTTTTTTTTTTTTGGGACAGTCTTGTCTGTGCCACAGGCTGAGTGA 257
y		181 CGGATCTTGGCTCACTGCACACTCTGCTCCACAGTTCAGAGAAATCTTCTTCCT 240
b		258 CGGATCTTGGCTCACTGCACACTCTGCTCCACAGTTCAGAGAAATCTTCTTCCT 317
y		241 TCCCAAGTAGCCAACAACGCCAGCTAAATTTTGTGTAATTTAGTAGAGCGGGTT 300
b		318 TCCCAAGTAGCCAACAACGCCAGCTAAATTTTGTGTAATTTAGTAGAGCGGGTT 377
y		301 GATGTTGTCAGCGTGTCTGTAACCTCTGACCTCAGTGTGATCCACCGCTCTCG 360
b		378 CATGTTGTCCAGCTGGTCTTGAACCTCTGACCTCAGTGTGATCCACCGCTCTCG 437
y		361 CAAAGTACTAGGATACAGGCGTCAGACCAACCGCTCGAGCGCCCTGGCGGTTTT 420
b		438 CAAAGTACTAGGATACAGGCGTGAGCCACCGGCTCGAGCGCCCTGGCGGTTTT 497
y		421 AGTAGAAAAGCTCATTATATACCACTTGCTTCNGTTGTCNNTTCAGTGAGAAC 480
b		498 AGTAGAAAAGCTCATTATATACCACTTGCTTCGTTGGC-TTCAGTGAGAACGA 556
y		481 AAATGCAAAATCNCCTTATTAGTTGTAGAAACAGATCTCAACAGCAGTTTGT 540
b		557 AAATGCAAAATCNCCTTATTAGTTGTAGAAACAGATCTCAACAGCAGTTTGT 616
y		541 ACCGAGGAAAAGTGGGAACTGTGCTCTGTGGCTTAGAGAAAGCGCGGTCGAC 600
b		617 ACCGAGGAAAAGTGGGAACTGTGCTGTGGCTTAGAGAAAGCGCGGTCGAC 676
y		601 TTCCCAAAAGGGCGAGTCTTCCCNCGCACCGCACCTGCTCAGAGTCCCGGGT 660
b		677 TTCCCAAAAGGGCGAGTCTTCCCNCGCACCGCACCTGCTCAGAGTCCCGGGT 736
y		661 AGAGCTCTCAGCTGTGGCCCTGGGCTCCGTTCTGTGTCCACACCGCTGGCTCT 720
b		737 AGAGCTCTCAGCTGTGGCCCTGGGCTCCGTTCTGTGTCCACACCGCTGGCTCT 796
y		721 CCCCTGGCGCACGCTCTCTAGAGCGGGGCGCGCGCACACCCGCGGAGGAGAG 780
b		797 CCCCTGGCGCACGCTCTCTAGAGCGGGGCGCGCGCACACCCGCGGAGGAGAG 856
y		781 GAGCGCGGAGCGCGCGGGAAGAGCGCGGGAAGGGGTCTGTCCACCGCGCTCT 840
b		857 GAGCGCGGAGCGCGCGGGAAGAGCGCGGGAAGGGGTCTGTCCACCGCGCTCT 916
y		841 CTTGCGTTCGGTCCCGCGCGCACTTTGGCTCTGCTCTGTCGCGCGCGCACTT 900
b		917 CTTGCGTTCGGTCCCGCGCGCACTTTGGCTCTGCTCTGTCGCGCGCGCACTT 976
y		901 CTTGCGTTCGGTCCCGCGCGCACTTTGGCTCTGCTCTGTCGCGCGCGCACTT 960
b		977 CTTGCGTTCGGTCCCGCGCGCACTTTGGCTCTGCTCTGTCGCGCGCGCACTT 1036
y		961 TTGCGCCCGCGCGCAAGAGAGCGGAGACCGAGCCGGTCTCCGCGCAGCGGAG 1020

Db	1037	TTTGCCCCCGCCGACAGGAGCGGGAGCCGCGAGCCGCGCTGCGCCGCGACGCGGGAGCTGCGAG	1096
QY	1021	TRACCTGGGGCAGATCCAAACACATCCTCCGCTGCGGGGCTCAGGAAGGACGACCGCAGCGGC	1080
Db	1097	TRACCTGGGGCAGATCCAAACACATCCTCCGCTGCGGGGCTCAGGAAGGACGACCGCAGCGGC	1156
QY	1081	ACCGGCAACCTGTGCGTATTGCGCATGCAAGCGCGCTACAGCCTGAGAGGTGACGCCGCG	1140
Db	1157	ACCGGCAACCTGTGCGTATTGCGCATGCAAGCGCGCTACAGCCTGAGAGGTGACGCCGCG	1216
QY	1141	GGCCCTGCGGGACGGGTGCGGGAAGGAGGGAGCGCGCTGGGGA	1187
Db	1217	GGCCCTGCGGGACGGGTGCGGGAAGGAGGGAGCGCGCTGGGGA	1263
RESULT 3			
ABEN95092	ID	ABEN95092 standard; DNA; 18596 BP.	
XX	AC	ABEN95092;	
AC	AC		
DT	13-AUG-2002	(first entry)	
DE	Gene #1590	used to diagnose liver cancer.	
XX	Gene;	liver cancer; ds; hepatocellular carcinoma; hepatotropic;	
KW	metastatic liver tumour;	cytostatic; expression profile; disease state;	
KW	disease progression;	drug toxicity; drug efficacy; drug metabolism.	
XX			
OS	Homo sapiens.		
XX			
FN	W0200229103-A2.		
XX			
PD	11-APR-2002.		
XX			
FF	02-OCT-2001; 2001WO-US30589.		
XX			
PR	02-OCT-2000; 2000US-237054P.		
XX	(GENE-) GENE LOGIC INC.		
PA			
XX	Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;		
PI	WPI; 2002-426119/45.		
XX			
DR			
XX			
PT	Diagnosing and detecting the progression of liver cancer.		
PT	hepatocellular carcinoma or metastatic liver tumor in a patient,		
PT	involves detecting the level of expression of two or more genes in a		
PT	liver tissue sample -		
XX			
PS	Claim 1; SEQ ID NO 1590; 298pp; English.		
XX			
CC	The invention relates to a novel method for diagnosing and detecting the		
CC	progression of liver cancer, hepatocellular carcinoma or metastatic liver		
CC	tumour in a patient, and differentiating metastatic liver cancer from		
CC	hepatocellular carcinoma in a patient, involving detecting the level of		
CC	expression of two or more genes represented in ABEN93503-ABEN97455 in a		
CC	tissue sample. The method of the invention has hepatotropic, and		
CC	cytostatic activity. The method is useful for diagnosing and detecting		
CC	the progression of liver cancer, hepatocellular carcinoma and metastatic		
CC	liver carcinoma in a patient. The method is useful for identifying		
CC	expression profiles which serve as useful diagnostic markers as well as		
CC	markers that can be used to monitor disease states, disease progression,		
CC	drug toxicity, drug efficacy and drug metabolism.		
CC	Note: The sequence data for this patent did not form part of the printed		
CC	specification, but was obtained in electronic format directly from WIPO		
CC	at ftp.wipo.int/pub/published_pat_sequences.		
XX			
SQ	Sequence 18596 BP; 4521 A; 3991 C; 4479 G; 5605 T; 0 other;		
Query Match 98.3%; Score 1167; DB 24; Length 18596;			
Best Local Similarity 99.2%; Pred. No. 1.1e-193;			
Matches 1179; Conservative 0; Mismatches 8; Indels 1; Gaps 1			

Query Match 98.3%; Score 1167; DB 24; Length 18596;
Best Local Similarity 99.2%; Prod. No. 1.1e-193;
Matches 1178; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

QY 1081 ACCGCGACCCCTGTCGGTATTTCGGCATGTCAGCGCGGCTACAGAGTGTGAGAGTGTGACCCGCG 1140
DB 1157 ACCGCGACCCCTGTCGGTATTTCGGCATGTCAGCGCGGCTACAGAGTGTGAGAGTGTGACCCGCG 1216
QY 1141 GCGCCCTGCGGAGCGGTGCGGGAAGAGGAGCGCGGCTGGGA 1187
DB 1217 GCGCCCTGCGGAGCGGTGCGGGAAGAGGAGCGCGGCTGGGA 1263

RESULT 4
ABK43334
ID ABK43334 standard; DNA; 18596 BP.
AC ABK43334;
XX
XX
DT 05-JUN-2002 (first entry)
XX
XX Human Thymidylate synthase gene sequence.
XX
XX HKNG1; ds; gene; chromosome 18p; bipolar affective disorder; BAD;
KW severe bipolar affective (mood) disorder; BP-I; schizophrenia; TS;
KW Hong Kong new gene 1; antimanic; antidepressant; neuroleptic;
KW Thymidylate synthase.
XX
XX Homo sapiens.
XX
XX WO200210366-A2.
XX
XX 07-FEB-2002.
XX
XX 02-AUG-2001; 2001WO-US24417.
XX
XX 02-AUG-2000; 2000US-0631275.
PR 28-NOV-2000; 2000US-0722544.
XX
XX (MILL-) MILLENNIUM PHARM INC.
PA (REGC) UNIV CALIFORNIA.
XX
XX Chen H, Freimer NB, Novak T;
PI WPI; 2002-195962/25.
XX
XX New nucleic acid molecule Hong Kong New Gene 1 (HKNG1), useful for
FT screening for molecules which modulate HKNG1 expression for the
PT treatment of bipolar disorder and schizophrenia -
XX Disclosure; Fig 44; 367pp; English.

The invention relates to an isolated nucleic acid molecule comprising a
nucleotide sequence that encodes a Hong Kong New Gene (HKNG) 1 gene
product. The human gene for HKNG1 is located on chromosome 18p in
an area associated with bipolar affective disorder, BAD. Also
included are an expression vector comprising the nucleic acid, a
host cell expressing the nucleic acid, an anti-HKNG1 antibody, a method
of identifying modulators of HKNG1, and identifying an individual (at
risk of) having HKNG1-mediated disorder comprising the presence
or absence of a polymorphism that correlates with an HKNG1 allele
associated with the disorder, where the presence of the polymorphism
indicates that the individual (is at risk of) having HKNG1-mediated
disorder. A (small molecule) compound which modulates (inhibits or
potentiates) expression of a HKNG1 gene or gene product in a human
individual is useful for the treatment of a HKNG1-mediated disorder
such as bipolar affective disorder (BAD), severe bipolar affective (mood)
disorder (BP-I) and schizophrenia. The present sequence is a
thymidylate synthase, TS, genomic DNA sequence. The gene for TS
overlaps that of HKNG1 and therefore TS may also be involved in the
diseases listed above.

Sequence 18596 BP; 4521 A; 3991 C; 4479 G; 5605 T; 0 other;
SQ

QY 1 GATCGCGCCTCCTCCTCAGCGCTGAGAGAGAGAGTCTCTGTCTCAAAAAA 60
DB 78 GATCGCGCCTCCTCCTCAGCGCTGAGAGAGAGAGTCTCTGTCTCAAAAAA 137
QY 61 AAAAAGACCCCGAGGCTCAAAACAAAAAAGCTCGGAGAGTCTCTCTTTT 120
DB 138 AAAAAGACCCCGAGGCTCAAAACAAAAAAGCTCGGAGAGTCTCTCTTTT 197
QY 121 TTTTCTTTTCTTTTCTTTTGGAGAGTCTCTCTCTGCGCCAGGCTGAGTACAAAGT 180
DB 198 TTTTCTTTTCTTTTCTTTTGGAGAGTCTCTCTCTGCGCCAGGCTGAGTACAAAGT 257
QY 181 CGGATCTTGGCTCAGTCAAGCTCTGCTCCAGGTTCAAGCAATTCCTCTGCTCAGCC 240
DB 258 CGGATCTTGGCTCAGTCAAGCTCTGCTCCAGGTTCAAGCAATTCCTCTGCTCAGCC 317
QY 241 TCCCAAGTAGCAGCCAGCCAGCTAAATTTTGTANTTTTAGTAGAGCGGGTTTCA 300
DB 318 TCCCAAGTAGCAGCCAGCCAGCTAAATTTTGTANTTTTAGTAGAGCGGGTTTCA 377
QY 301 CATGTTGCCAGGCTGCTCTGAGCTCTGAGCTCAGTCAATCCACCCGCTCGGCCCCC 360
DB 378 CATGTTGCCAGGCTGCTCTGAGCTCTGAGCTCAGTCAATCCACCCGCTCGGCCCCC 437
QY 361 CAAAGTACTAGGATTAAGCGCTGAGCAACCGCTCAGCGCTCTGGCGTTTAAATCA 420
DB 438 CAAAGTACTAGGATTAAGCGCTGAGCAACCGCTCAGCGCTCTGGCGTTTAAATCA 497
QY 421 AGTAGAAAAGTGCATTAATCACTGCTTCTGTTGCTTCACTGAGAGCAAGAAATGG 480
DB 498 ACTAGAAAAGTGCATTAATCACTGCTTCTGTTGCTTCACTGAGAGCAAGAAATGG 556
QY 481 AAATGCAATCTTATTTAGTTCTAGGAAACAGATCTCAACAGCAGTCTTCTGACAG 540
DB 557 AAATGCAATCTTATTTAGTTCTAGGAAACAGATCTCAACAGCAGTCTTCTGACAG 616
QY 541 ACCGAGAGAAACCTGGGAACTGTGCTGCTTGTAGAGAGCGCGCTCGACAGAGCG 600
DB 617 ACCGAGAGAAACCTGGGAACTGTGCTGCTTGTAGAGAGCGCGCTCGACAGAGCG 676
QY 601 TTCCCAAGAGCGCAGTCTTCCCGCCAGCCAGCTCCTGCTGAGAGTCTCGGCTTCT 660
DB 677 TTCCCAAGAGCGCAGTCTTCCCGCCAGCCAGCTCCTGCTGAGAGTCTCGGCTTCT 736
QY 661 AAGACTCTCAGTCTGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
DB 737 AAGACTCTCAGTCTGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 796
QY 721 CCGCTGCGCAGCTCTCTAGAGGGGGCGCGCGACCCCGCGAGAGAGAGAGCG 780
DB 797 CCGCTGCGCAGCTCTCTAGAGGGGGCGCGCGACCCCGCGAGAGAGAGAGCG 856
QY 781 GAGCGGAGAGCGCGCGGAGAGAGCGCGGAGAGAGAGAGAGAGAGAGAGAGAG 840
DB 857 GAGCGGAGAGCGCGCGGAGAGAGAGCGCGGAGAGAGAGAGAGAGAGAGAGAG 916
QY 841 CTTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
DB 917 CTTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 976
QY 901 CTTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
DB 977 CTTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1036
QY 961 TTGCCCCCGCGCAG 1020
DB 1037 TTGCCCCCGCGCAG 1096
QY 1021 TACTTGGGCGAGATCAACACATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
DB 1097 TACTTGGGCGAGATCAACACATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1156

Query Match 98.3%; Score 1167; DB 24; Length 18596;
Best Local Similarity 99.2%; Pred. No. 1.1e-193;
Matches 1178; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

Y 1 GATCGCGCACTGACCTCCAGCCTGGGTGAGAGAGGAGACTCTGTCTCAAAAAA 60
b 78 GATCGCGCACTGACCTCCAGCCTGGGTGAGAGAGGAGACTCTGTCTCAAAAAA 137
Y 61 AAAAAAGACCGCGAGGCTCAAAACAAAAAAACCTCGAAAAAGCCCTGGGGCTCTTTT 120
b 138 AAAAAAGACCGCGAGGCTCAAAACAAAAAAACCTCGAAAAAGCCCTGGGGCTCTTTT 197
Y 121 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 180
b 198 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 257
Y 181 CGATCTTGGCTCACTGCAACCTCTGCTCCAGGTTCAAGCAATCTTCTGCTCAGCC 240
b 258 CGATCTTGGCTCACTGCAACCTCTGCTCCAGGTTCAAGCAATCTTCTGCTCAGCC 317
Y 241 TCCCAAGTAGCCACACAGCCAGCTAATTTTGTANTTTTAGTAGAGAGGGGTTCAC 300
b 318 TCCCAAGTAGCCACACAGCCAGCTAATTTTGTACTTTTAGTAGAGAGGGGTTCAC 377
Y 301 CATGTTGTCAGGCTGCTCTGAACTCTGACCTCAGGTGATCCACCCGCTCGGCCCCC 360
b 378 CATGTTGTCAGGCTGCTCTGAACTCTGACCTCAGGTGATCCACCCGCTCGGCCCCC 437
Y 361 CAAAGTACTAGGATTAAGCGGTGAGCCACCGCTCAGCGCTCGGGTTTTTAATCA 420
b 438 CAAAGTACTAGGATTAAGCGGTGAGCCACCGCTCAGCGCTCGGGTTTTTAATCA 497
Y 421 AGTAGAAAAGCTGATTAATACACTTGTCTGTTGTTGNTTCACTGAGAGAGAAATGG 480
b 498 AGTAGAAAAGCTGATTAATACACTTGTCTGTTGTTGNTTCACTGAGAGAGAAATGG 556
Y 481 AAATGCAATCNCCTTATTTAGTTGAGGAAACAGATCTCAACAGCAGTTTGTGACAG 540
b 557 AAATGCAATCNCCTTATTTAGTTGAGGAAACAGATCTCAACAGCAGTTTGTGACAG 616
Y 541 ACCGAGGAAACCTGGAATCTGTGCTGTGCTTAGAGAGAGCGGTCGACACAGCG 600
b 617 ACCGAGGAAACCTGGAATCTGTGCTGTGCTTAGAGAGAGCGGTCGACACAGCG 676
Y 601 TTCCCAAGAGCGGAGCTCTTCCGCGCACCGCTGCTGCTGCTGCTGCTGCTGCT 660
b 677 TTCCCAAGAGCGGAGCTCTTCCGCGCACCGCTGCTGCTGCTGCTGCTGCTGCT 736
Y 661 AAGAATCTCAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
b 737 AAGAATCTCAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 796
Y 721 CCGTGGGCACTGCTCTTAGAGCGGGGGCCCGCGACCCCGCGAGCAGGAGAGGCG 780
b 797 CCGTGGGCACTGCTCTTAGAGCGGGGGCCCGCGACCCCGCGAGCAGGAGAGGCG 856
Y 781 GAGCGCGGAGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
b 857 GAGCGCGGAGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 916
Y 841 CTTGCTCTCGCTCGCGCGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
b 917 CTTGCTCTCGCTCGCGCGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 976
Y 901 CTTGCTCTCGCTCGCGCGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
b 977 CTTGCTCTCGCTCGCGCGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1036
Y 961 TTGCCCCCGCGCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1020
b 1037 TTGCCCCCGCGCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1096
Y 1021 TACCTGGGAGGAGTCAACACATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
b 1097 TACCTGGGAGGAGTCAACACATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1156

QY 1081 ACCGCAACCTGCTGCTATTTGGCATGCGAGCGGCTACAGCTGAGAGGTGACGCGG 1140
Db |||||
Db 1157 ACCGCAACCTGCTGCTATTTGGCATGCGAGCGGCTACAGCTGAGAGGTGACGCGG 1216
QY 1141 GSCCCTGCGGACGCGGTGGCGGGAAGAGGAGGCGCGGCTTGGGA 1187
Db |||||
Db 1217 GSCCCTGCGGACGCGGTGGCGGGAAGAGGAGGCGCGGCTTGGGA 1263
RESULT 5
ABL62854
ID ABL62854 standard; DNA; 18596 BP.
XX AC ABL62854;
XX DT 15-MAY-2002 (first entry)
XX
DE Breast cancer related gene sequence SEQ ID NO:1191.
XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumor; cancerous;
KW cytosolic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200194629-A2.
XX PD 13-DEC-2001.
XX PF 30-MAY-2001; 2001WO-US10838.
XX PR 05-JUN-2000; 2000US-209473P.
PR 05-JUN-2000; 2000US-209531P.
PR 18-SEP-2000; 2000US-233133P.
PR 18-SEP-2000; 2000US-233617P.
PR 20-SEP-2000; 2000US-234009P.
PR 20-SEP-2000; 2000US-234034P.
PR 20-SEP-2000; 2000US-234052P.
PR 22-SEP-2000; 2000US-234509P.
PR 22-SEP-2000; 2000US-234567P.
PR 25-SEP-2000; 2000US-234923P.
PR 25-SEP-2000; 2000US-234924P.
PR 25-SEP-2000; 2000US-235077P.
PR 25-SEP-2000; 2000US-235134P.
PR 25-SEP-2000; 2000US-235280P.
PR 26-SEP-2000; 2000US-235637P.
PR 26-SEP-2000; 2000US-235638P.
PR 27-SEP-2000; 2000US-235711P.
PR 27-SEP-2000; 2000US-235720P.
PR 27-SEP-2000; 2000US-235840P.
PR 27-SEP-2000; 2000US-235863P.
PR 28-SEP-2000; 2000US-236028P.
PR 28-SEP-2000; 2000US-236032P.
PR 28-SEP-2000; 2000US-236033P.
PR 28-SEP-2000; 2000US-236034P.
PR 28-SEP-2000; 2000US-236109P.
PR 28-SEP-2000; 2000US-236111P.
PR 29-SEP-2000; 2000US-236842P.
PR 29-SEP-2000; 2000US-236891P.
PR 02-OCT-2000; 2000US-237172P.
PR 02-OCT-2000; 2000US-237173P.
PR 02-OCT-2000; 2000US-237278P.
PR 02-OCT-2000; 2000US-237294P.
PR 02-OCT-2000; 2000US-237295P.
PR 02-OCT-2000; 2000US-237316P.
PR 03-OCT-2000; 2000US-237435P.
PR 03-OCT-2000; 2000US-237596P.
PR 03-OCT-2000; 2000US-237604P.
PR 03-OCT-2000; 2000US-237606P.
PR 03-OCT-2000; 2000US-237608P.
PR 01-NOV-2000; 2000US-244867P.

PR 01-NOV-2000; 2000US-245084P.
XX (AVAL-) AVALON PHARM.
XX
XX
PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX
XX WPI; 2002-188264/24.
XX
XX
XX Screening for anti-neoplastic agent involves exposing cells to a
PT chemical agent to be tested for anti-neoplastic activity, and
PT determining a change in expression of a gene of a signature gene set
-
XX
XX Claim 1; SEQ ID 1191; 44pp; English.
XX
XX The present invention describes a method (M1) for screening for an
CC anti-neoplastic agent. The method involves exposing cells to a chemical
CC agent to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening
CC an anti-neoplastic agent, and can be used for producing a product which
CC is the data collected with respect to the anti-neoplastic agent as a
CC result of M1, and the data is sufficient to convey the chemical
CC structure and/or properties of the agent. M1 can be used in the
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC carcinoma, papillary carcinoma and Wilm's tumour.
XX
XX
SQ Sequence 18596 BP; 4521 A; 3991 C; 4479 G; 5605 T; 0 other;
Query Match 98.3%; Score 1167; DB 24; Length 18596;
Best Local Similarity 99.2%; Pred. No. 1.le-193;
Matches 1178; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

QY 1 GATCGCGCACTGCACTCCAGCTGGGTGAGAGAGCGAGACTCTGCTCAAAAAA 60
DB 78 GATCGCGCACTGCACTCCAGCTGGGTGAGAGAGCGAGACTCTGCTCAAAAAA 137
QY 61 AAAAAAGCGCGAGCGGTCAAAACAAAAACCTCGGAAAGCCCTGGGGTCTTTTTTT 120
DB 138 AAAAAAGCGCGAGCGGTCAAAACAAAAACCTCGGAAAGCCCTGGGGTCTTTTTTT 197
QY 121 TTTTTTTTTTTTTTTTTTTTGGGACAGTCTTGCTGTGTCGCCAGGTGGAGTACAAATGGT 180
DB 198 TTTTTTTTTTTTTTTTTTTTGGGACAGTCTTGCTGTGTCGCCAGGTGGAGTACAAATGGT 257
QY 181 CGGATCTTGGGTCTCACTCAACCTCTGCCCTCCAGGTTCAGCAATCTTCTGCTCAGCC 240
DB 258 CGGATCTTGGGTCTCACTCAACCTCTGCCCTCCAGGTTCAGCAATCTTCTGCTCAGCC 317
QY 241 TCCCAAGTAGCCACAGCCCGCAGCTAATTTTGTANTTTTGTAGAGAGCGGGTTTTCAC 300
DB 318 TCCCAAGTAGCCACAGCCCGCAGCTAATTTTGTACTTTTGTAGTAGAGCGGGTTTTCAC 377
QY 301 CATGTGTCCAGGCTGGTCTGCTGACCTCAGGTGATCCACCGGCTCGGCGCCCC 360
DB 378 CATGTGTCCAGGCTGGTCTGAACTCCTGACTCAGGTGATCCACCGGCTCGGCCCC 437
QY 361 CAAAGTACTAGGATTAAGGCGTGAGCCACCGCGTCCAGCGCTCGGCGTTTAAATCA 420
DB 438 CAAAGTACTAGGATTAAGGCGTGAGCCACCGGCTCCAGCGCTCGGCGTTTAAATCA 497
QY 421 AGTAGAAAAAGCTGCATTATACCACTTGCTTCNGTTTCNNTTCAGTGAGAACGAAATGG 480
DB 498 AGTAGAAAAAGCTGCATTATACCACTTGCTTCGTTGCTTTCAGTGAGAACGAAATGG 556
QY 481 AAATGCAATCNCTTATTAGTTGTAGGAAACAGATCTCAACACAGCTTTTGTGNGACAAG 540

DB 557 AAATGCAATCCCTTATTAGTTGTAGGAAACAGATCTCAACACAGCAGCTTTTGTGACAAG 616
QY 541 ACGCAGCAAAAACGTGGGAACCTGCTGCTGTGCTTGTAGAGAGGCGGCTGACACGAGCG 600
DB 617 ACGCAGGAAACCTGGGAACCTGCTGCTGTGCTTGTAGAGAGGCGGCTGACACGAGCG 676
QY 601 TTCCCAAAAGGCGCAGTCTTCCNGCCACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
DB 677 TTCCCAAAAGGCGCAGTCTTCCAGCACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 736
QY 661 AAGACTCTCAGCTGTGGCGCTTCCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
DB 737 AAGACTCTCAGCTGTGGCGCTTCCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 796
QY 721 CCGCTGGCGCACTCTTCTAGAGCGGGCGCGCGCGACCCCGCGCAGCAGGAGAGCGG 780
DB 797 CCGCTGGCGCACTCTCTAGAGCGGGCGCGCGCGACCCCGCGCAGCAGGAGAGCGG 856
QY 781 GAGCGGGGAGCGCGCGCGGGAAGGCGCGGAAAGGGGCTCTGCTGCTGCTGCTGCTGCTGCT 840
DB 857 GAGCGCGGAGCGCGCGCGGGAAGGCGCGGAAAGGGGCTCTGCTGCTGCTGCTGCTGCTGCT 916
QY 841 CTTGCTCTCGCTCCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
DB 917 CTTGCTCTCGCTCCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 976
QY 901 CTTGCTCTCGCTCCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
DB 977 CTTGCTCTCGCTCCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1036
QY 961 TTGCCCCCGCGCGCAGCAGGAGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
DB 1037 TTGCCCCCGCGCGCAGCAGGAGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1096
QY 1021 TACCTGGGCGAGATCCAAACACATCTTCCGCTGCGCGCTGAGGAGGAGCAGCAGCGGCG 1080
DB 1097 TACCTGGGCGAGATCCAAACACATCTTCCGCTGCGCGCTGAGGAGGAGCAGCAGCGGCG 1156
QY 1081 ACGGCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
DB 1157 ACGGCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1216
QY 1141 GCGCCCTCGGAGAGCGGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1187
DB 1217 GCGCCCTCGGAGAGCGGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1263

RESULT 6
ABL63078 standard; DNA; 18596 BP.
ID ABL63078
XX
XX ABL63078;
XX
XX 15-MAY-2002 (first entry)
XX
XX
XX Breast cancer related gene sequence SEQ ID NO:1415.
XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
XX stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
XX cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
XX gene; da.
XX Homo sapiens.
XX
XX WO200194629-A2.
XX
XX 13-DEC-2001.
XX
XX 30-MAY-2001; 2001WO-US10838.
XX
XX 05-JUN-2000; 2000US-209473P.
XX
XX 05-JUN-2000; 2000US-209531P.
XX
XX 18-SEP-2000; 2000US-233133P.
XX

R 18-SEP-2000; 2000US-233617P.
R 20-SEP-2000; 2000US-234009P.
R 20-SEP-2000; 2000US-234034P.
R 20-SEP-2000; 2000US-234052P.
R 22-SEP-2000; 2000US-234509P.
R 22-SEP-2000; 2000US-234567P.
R 25-SEP-2000; 2000US-234923P.
R 25-SEP-2000; 2000US-234924P.
R 25-SEP-2000; 2000US-235077P.
R 25-SEP-2000; 2000US-235082P.
R 25-SEP-2000; 2000US-235134P.
R 25-SEP-2000; 2000US-235280P.
R 26-SEP-2000; 2000US-235637P.
R 26-SEP-2000; 2000US-235638P.
R 27-SEP-2000; 2000US-235711P.
R 27-SEP-2000; 2000US-235720P.
R 27-SEP-2000; 2000US-235840P.
R 27-SEP-2000; 2000US-235863P.
R 28-SEP-2000; 2000US-236028P.
R 28-SEP-2000; 2000US-236032P.
R 28-SEP-2000; 2000US-236033P.
R 28-SEP-2000; 2000US-236034P.
R 28-SEP-2000; 2000US-236109P.
R 28-SEP-2000; 2000US-236111P.
R 28-SEP-2000; 2000US-236112P.
R 29-SEP-2000; 2000US-236842P.
R 29-SEP-2000; 2000US-236891P.
R 02-OCT-2000; 2000US-237172P.
R 02-OCT-2000; 2000US-237173P.
R 02-OCT-2000; 2000US-237278P.
R 02-OCT-2000; 2000US-237294P.
R 02-OCT-2000; 2000US-237295P.
R 02-OCT-2000; 2000US-237316P.
R 03-OCT-2000; 2000US-237425P.
R 03-OCT-2000; 2000US-237598P.
R 03-OCT-2000; 2000US-237604P.
R 03-OCT-2000; 2000US-237606P.
R 03-OCT-2000; 2000US-237608P.
R 01-NOV-2000; 2000US-244867P.
R 01-NOV-2000; 2000US-245084P.
(AVAL-) AVALON PHARM.
Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
Soppet DR, Weaver Z;
WPI; 2002-188264/24.
Screening for anti-neoplastic agent involves exposing cells to a
chemical agent to be tested for anti-neoplastic activity, and
determining a change in expression of a gene of a signature gene set -
Claim 1; SEQ ID 1415; 44pp; English.
The present invention describes a method (M1) for screening for an
anti-neoplastic agent. The method involves exposing cells to a chemical
agent to be tested for anti-neoplastic activity. Determining a change in
expression of at least one gene (I) of a signature gene set, where (I)
comprises a sequence (S) selected from 847 sequences (given in AB51664
to AB70110), or is at least 95% identical to (S), where a change in
expression is indicative of anti-neoplastic activity. (I) has cytostatic
activity and can be used in gene therapy. M1 can be used for screening
an anti-neoplastic agent, and can be used for producing a product which
is the data collected with respect to the anti-neoplastic agent as a
result of M1, and the data is sufficient to convey the chemical
structure and/or properties of the agent. M1 can be used in the
treatment of cancer such as colon, breast, stomach, lung, thyroid,
oesophageal, ovarian, kidney, prostate or pancreatic cancer,
adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
carcinoma, Papillary carcinoma and Wilms' tumour.
Sequence 18596 BP; 4521 A; 3991 C; 4479 G; 5605 T; 0 other;

Query Match 98.3%; Score 1167; DB 24; Length 18596;
Best Local Similarity 99.2%; Pred. No. 1.1e-193;
Matches 1178; Conservative 0; Mismatches 8; Indels 1; Gaps 1;
QY 1 GATCGCCCACTGCACTCCAGCCTGGGTGAGAGCGGAGACTCTGTCTCAAAAAA 60
DB 78 GATCGCCCACTGCACTCCAGCCTGGGTGAGAGCGGAGACTCTGTCTCAAAAAA 137
QY 61 AAAAGAGCCGCGAGGCTCAAAACAAAACCTCGGAAAGCCCTGGCGGTCTTTT 120
DB 138 AAAAGAGCCGCGAGGCTCAAAACAAAACCTCGGAAAGCCCTGGCGGTCTTTT 197
QY 121 TTTTCTTTTCTTTTCTTTTGGGACAGCTTCTCTCTCGCCAGCTCGAGTACAATG 180
DB 198 TTTTCTTTTCTTTTCTTTTGGGACAGCTTCTCTCTCGCCAGCTCGAGTACAATG 257
QY 181 CGGATCTTGGCTCACTGCAACCTCTGCTCTCCAGGTTCAAGCAATCTTCTGCTCAGCC 240
DB 258 CGGATCTTGGCTCACTGCAACCTCTGCTCTCCAGGTTCAAGCAATCTTCTGCTCAGCC 317
QY 241 TCCCAAGTAGCCACACGCGCCAGCTAAATTTTGTATTTTATAGAGACGGGGTTTCA 300
DB 318 TCCCAAGTAGCCACACGCGCCAGCTAAATTTTGTATTTTATAGAGACGGGGTTTCA 377
QY 301 CATGTTCTCCAGGCTGCTTNGAACTCTGACCTCAGGTGATCCACCGCTCGGCCCC 360
DB 378 CATGTTCTCCAGGCTGCTTNGAACTCTGACCTCAGGTGATCCACCGCTCGGCCCC 437
QY 361 CAAGTACTAGGATTACAGGCTGAGCCACCGCTCCAGCGCCCTGGCGGTCTTTTAA 420
DB 438 CAAGTACTAGGATTACAGGCTGAGCCACCGCTCCAGCGCCCTGGCGGTCTTTTAA 497
QY 421 AGTAGAAAAGCTGATTAACATCTGCTTTCTGCTTNGTTCAGTGAGAGCAAGAAATGG 480
DB 498 AGTAGAAAAGCTGATTAACATCTGCTTTCTGCTTNGTTCAGTGAGAGCAAGAAATGG 556
QY 481 AAATGCAATCTTATAGTGTAGAAACAGATCTCAACAGCAGTTTGTGTGCAAG 540
DB 557 AAATGCAATCTTATAGTGTAGAAACAGATCTCAACAGCAGTTTGTGTGCAAG 616
QY 541 ACCGCGAGAAACGTTGGGAACTGTGTGCTGTAGAGAGGCGCGTGCACAGACGG 600
DB 617 ACCGCGAGAAACGTTGGGAACTGTGTGCTGTAGAGAGGCGCGTGCACAGACGG 676
QY 601 TTCCAAAAGGGCGCAGTCTTCCGNGCCACCGCACCTGCTCCAGGTTCCCGGTTTCT 660
DB 677 TTCCAAAAGGGCGCAGTCTTCCGNGCCACCGCACCTGCTCCAGGTTCCCGGTTTCT 736
QY 661 AAGACTCTCAGTGTGGCCCTGGGCTCCGTTCTGTGCAACACCGGTGCTCTGCGTTTC 720
DB 737 AAGACTCTCAGTGTGGCCCTGGGCTCCGTTCTGTGCAACACCGGTGCTCTGCGTTTC 796
QY 721 CCCCTGGCGCAGCTCTCTAGAGCGGGGCGCGCGACCCCGCGAGCAGGAGGGCG 780
DB 797 CCCCTGGCGCAGCTCTCTAGAGCGGGGCGCGCGACCCCGCGAGCAGGAGGGCG 856
QY 781 GAGCGCGGAGCGCGCGGGAAGAGGCGCGGAGGGGTCTGCAACCGCGCACCTTGG 840
DB 857 GAGCGCGGAGCGCGCGGGAAGAGGCGCGGAGGGGTCTGCAACCGCGCACCTTGG 916
QY 841 CCTGCTCTCCGTCGCGCGGCGCAGCTTGGCTGCTGCTCCGTCGCGCGCGCACCTTGCCTG 900
DB 917 CCTGCTCTCCGTCGCGCGGCGCAGCTTGGCTGCTGCTCCGTCGCGCGCGCACCTTGCCTG 976
QY 901 CCTCGCTCCCGCGCGCGCGCAGCTGCTGCTGCGCGCTCGAGCTGCGCGCGCGCGCC 960
DB 977 CCTCGCTCCCGCGCGCGCGCAGCTGCTGCTGCGCGCTCGAGCTGCGCGCGCGCGCC 1036
QY 961 TTGCCCCCGCGCGCAGAGCGGAGCGCGAGCTGCTGCTGCGCGCAGGAGCTGCGAG 1020
DB 1037 TTGCCCCCGCGCGCAGAGCGGAGCGCGAGCTGCTGCTGCGCGCAGGAGCTGCGAG 1096
QY 1021 TACCTGGGCGAGATCCAAACATCTCTCGCTGCGGCTCAGGAGGAGCAGCAGCGGC 1080

Db 314 A 314

RESULT 11

AAS94945

ID AAS94945 standard; DNA; 3298 BP.

XX AAS94945;

XX 14-FEB-2002 (first entry)

XX Human DNA sequence #200 expressed during foam cell differentiation.

XX Human; foam cell differentiation; atherosclerosis; cerebral stroke;

XX cardiovascular disorder; coronary artery disease; gene therapy; ds.

XX Homo sapiens.

XX WO200177389-A2.

XX 18-OCT-2001.

XX 04-APR-2001; 2001WO-US11128.

XX 05-APR-2000; 2000US-195106P.

XX (INCY-) INCYTE GENOMICS INC.

XX Shiffman D, Somogyi R, Lawn R, Seilhamer JJ, Porter GJ, Mikita T;

PI Tai J;

XX WPI; 2002-010925/01.

XX Composition useful for diagnosis of conditions, disorders or diseases

PT associated with atherosclerosis, comprises several polynucleotides that

PT are differentially expressed in foam cell development -

XX Claim 1; Page 250-251; 315pp; English.

XX The present invention relates to the isolation of human polynucleotide

CC sequences that are differentially expressed during foam cell

CC differentiation. The polynucleotide sequences of the invention or a

CC composition comprising these polynucleotides are useful as a high

CC throughput method for detecting altered expression of one or more

CC polynucleotides in a sample. The polynucleotides can be used in the

CC diagnosis of disorders associated with foam cell development such as

CC atherosclerosis, cerebral stroke, and cardiovascular disorders such as

CC coronary artery disease. The polynucleotide sequences can also be used

CC as PCR primers and probes. The polynucleotides of the invention are also

CC useful in gene therapy. AAS94746-AAS95021 represent the human

CC polynucleotide sequences of the invention which are differentially

CC expressed during foam cell differentiation.

XX SQ Sequence 3298 BP; 898 A; 709 C; 744 G; 919 T; 28 other;

Query Match 24.3%; Score 288.8; DB 24; Length 3298;

Best Local Similarity 99.0%; Pred. No. 1.8e-41;

Matches 290; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 841 CTTGCTCTCCCGCGCGCCACTTGGCTGCTCCGTCCTCCGCGCCACTTGGCTG 900

Db 1 CTTGCTCTCCGTCGCGCGCCACTTGGCTGCTCCGTCCTCCGCGCCACTTGGCTG 60

QY 901 CTTGCTCTCCCGCGCGCCACTTGGCTGCTCCGTCCTCCGCGCCACTTGGCTG 960

Db 61 CTTGCTCTCCCGCGCGCCACTTGGCTGCTCCGTCCTCCGCGCCACTTGGCTG 120.

QY 961 TTGCTCTCCCGCGCGCCACTTGGCTGCTCCGTCCTCCGCGCCACTTGGCTG 1020

Db 121 TTGCTCTCCCGCGCGCCACTTGGCTGCTCCGTCCTCCGCGCCACTTGGCTG 180

QY 1021 TTACCTGGGCGAGATCCACACATCTTCCTGCTGCGGCGTTCAGGAAGGACGCGCGC 1080

Db 181 TACCTGGGCGAGATCCACACATCTTCCTGCTGCGGCTCAGGAAGGACGCGCGCGGC 240

QY 1081 ACCGGCACCTCTCGGTATTCGGCATGCGAGCGCGCTACAGCTGAGAGTGA 1133

Db 241 ACCGGCACCTCTCGGTATTCGGCATGCGAGCGCGCTACAGCTGAGAGTGA 293

RESULT 12

ABQ39526

ID ABQ39526 standard; DNA; 764 BP.

XX ABQ39526;

XX 12-JUL-2002 (first entry)

XX Oligonucleotide for detecting cytosine methylation SEQ ID NO 26117.

XX Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;

XX drug; side effect; cancer; central nervous system; cardiovascular;

XX gastrointestinal; respiratory system; single nucleotide polymorphism;

XX SNP; cell differentiation; ds.

XX Homo sapiens.

XX WO200218632-A2.

XX 07-MAR-2002.

XX 01-SEP-2001; 2001WO-EF10074.

XX 01-SEP-2000; 2000DE-1043826.

XX 05-SEP-2000; 2000DE-1044543.

XX (EPIC-) EPIGENOMICS AG.

XX Olek A, Piepenbrock C, Berlin K, Guetig D;

XX WPI; 2002-371829/40.

XX Determining the degree of cytosine methylation in genomic DNA, useful

PT for diagnosis and prognosis, comprises selective hybridization of

PT amplicons from chemically treated DNA -

XX Claim 12; 56pp + Sequence Listing; 56pp; German.

XX This invention describes a novel method for determining the degree of

CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a

CC genomic sample of DNA. The sample is treated chemically to convert

CC cytosine (C) but not methylated C, to uracil, then part of the genomic

CC DNA that contains the target C is amplified to form a labeled amplicon.

CC The amplicon is hybridised to two classes, each with at least one

CC member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers

CC and the degree of hybridisation to both classes is determined from the

CC label on the amplicon. From the ratio of labels hybridised to the two

CC classes of oligomers, the degree of methylation is calculated. The method

CC is used: (i) for diagnosis and/or prognosis of side effects of

CC therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders

CC of the central nervous, cardiovascular, gastrointestinal and respiratory

CC systems etc., particularly by detecting mutations or single nucleotide

CC polymorphisms (SNP's); and (ii) for differentiation of cell or tissue

CC types and for investigating cell differentiation. The method allows the

CC methylation status of many C residues to be determined simultaneously.

CC ABQ3410-ABQ34121 represent genomic DNA sequences used to illustrate the

CC method for determining the degree of cytosine methylation described in

CC the disclosure of the invention.

XX SQ Sequence 764 BP; 109 A; 98 C; 283 G; 274 T; 0 other;

Query Match 19.1%; Score 226.6; DB 24; Length 764;

Best Local Similarity 71.8%; Pred. No. 1e-30;

Matches 324; Conservative 0; Mismatches 99; Indels 28; Gaps 1;

QY 737 TTACGAGCGGGGGCGCGCGACCCCGCGAGAGGAGGCGGAGCGGCGCGCGCG 796

b 1 TTTAGAGCGGGGCTCTCGGATTCCTCGATAGGAGGCGGAGCGGAGCGGTG 60
 y 797 CCGGAAAAGCGCGCGAAGGGTCTCTGCAACCGGCGCACTTGCGCTCGCTCCGTCGCCG 856
 b 61 CCGGAAAAGCGCGCGAAGGGTCTCTGCAACCGGCGCACTTGCGCTCGCTCCGTCGCCG 92
 y 857 CCGGCACTTGGCTGCTCTCGCTCGCGCGCGCGCACTTGCGCTCGCTCCGTCGCCG 916
 b 93 CCGGCACTTGGCTGCTCTCGCTCGCGCGCGCGCACTTGCGCTCGCTCCGTCGCCG 152
 y 917 GCGCGCGCATCGCTGTCGCGCTCGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCG 976
 b 153 GTCGCGTAATGTTTGCGTCTCGGAGTGTGCGGTCGCTGCTGCTGCTGCTGCTGCT 212
 y 977 AGGAGCGGAGCGCGCGCGCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1036
 b 213 AGGAGCGGAGCGCGCGCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 272
 y 1037 AACACATCTCTGCTGCGCGCTCAGAGAGGAGCGCGCGCGCGCGCGCGCGCGCG 1096
 b 273 AATAATATTTTCTGTTGCGCGCTTAGAGAGGAGCGCGCGCGCGCGCGCGCGCG 332
 y 1097 TATTCGCGATCGAGCGCGCTACGCTCAGAGGCTGACGCGCGCGCGCGCGCGCG 1156
 b 333 TATTCGCGATCGAGCGCGCTTAGAGGAGGAGCGCGCGCGCGCGCGCGCGCGCG 392
 y 1157 GTGCGCGGAGGAGGAGGCGCGCTGCGGGA 1187
 b 393 GTGCGCGGAGGAGGAGGCGCGCTGCGGGA 423

RESULT 13
 ABQ39527/c
 ID ABQ39527 standard; DNA; 764 BP.
 XX AC ABQ39527;
 XX DT 12-JUL-2002 (first entry)
 XX OLigonucleotide for detecting cytosine methylation SEQ ID NO 26119.
 XX Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
 XX drug; side effect; cancer; central nervous system; cardiovascular;
 XX gastrointestinal; respiratory system; single nucleotide polymorphism;
 XX SNP; cell differentiation; ds.
 XX OS Homo sapiens.
 XX WO200218632-A2.
 XX PD 07-MAR-2002.
 XX PF 01-SEP-2001; 2001WO-EP10074.
 XX PR 01-SEP-2000; 2000DE-1043826.
 XX PR 05-SEP-2000; 2000DE-1044543.
 XX PA (EPIG-) EPIGNOMICS AG.
 XX PI Olek A, Piepenbrock C, Berlin K, Guetig D;
 XX WPI; 2002-371829/40.
 XX Determining the degree of cytosine methylation in genomic DNA, useful
 XX for diagnosis and prognosis, comprises selective hybridization of
 XX amplicons from chemically treated DNA -
 XX Claim 12; 56pp + Sequence Listing; 56pp; German.
 XX This invention describes a novel method for determining the degree of
 XX methylation of a particular cytosine in a motif 5'-CpG-3', present in a
 XX genomic sample of DNA. The sample is treated chemically to convert

CC cytosine (C) but not methylated C, to uracil, then part of the genomic
 CC DNA that contains the target C is amplified to form a labeled amplicon.
 CC The amplicon is hybridised to two classes, each with at least one
 CC member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers
 CC and the degree of hybridisation to both classes is determined from the
 CC label on the amplicon. From the ratio of labels hybridised to the two
 CC classes of oligomers, the degree of methylation is calculated. The method
 CC is used: (i) for diagnosis and/or prognosis of side effects of
 CC therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders
 CC of the central nervous, cardiovascular, gastrointestinal and respiratory
 CC systems etc., particularly by detecting mutations or single nucleotide
 CC polymorphisms (SNP/s); and (ii) for differentiation of cell or tissue
 CC types and for investigating cell differentiation. The method allows the
 CC methylation status of many C residues to be determined simultaneously.
 CC ABQ3410-ABQ54121 represent genomic DNA sequences used to illustrate the
 CC method for determining the degree of cytosine methylation described in
 CC the disclosure of the invention.
 XX Sequence 764 BP; 274 A; 283 C; 98 G; 109 T; 0 other;

Query Match 19.1%; Score 226.6; DB 24; Length 764;
 Best Local Similarity 71.8%; Pred. No. 18-30; Indels 28; Gaps 1;
 Matches 324; Conservative 0; Mismatches 99;
 QY 737 TCTAGAGCGGGGCG 796
 Db 764 TTTAGAGCGGGGCG 705
 QY 797 CCGGAAAAGCG 856
 Db 704 CCGGAAAAGCG 673
 QY 857 CGCGCGCATGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 916
 Db 672 CGCGGTATTTGGTTCGTTTCGTTTCGTTTCGTTTCGTTTCGTTTCGTTTCGTT 613
 QY 917 GCGCGCGCATGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 976
 Db 612 CTCGCGTATATGTTTGGTTCGTTTCGTTTCGTTTCGTTTCGTTTCGTTTCGTT 553
 QY 977 AGGAGCGGAGCG 1036
 Db 552 AGGAGCGGAGCG 493
 QY 1037 AACACATCTCTGCTGCGCGCTCAGAGAGGAGCGCGCGCGCGCGCGCGCGCGCG 1096
 Db 492 AATATATTTTTCGTTTCGCGCGCTTAGAGAGGAGCGCGCGCGCGCGCGCGCG 433
 QY 1097 TATTCGCGATCGAGCGCGCTACGCTCAGAGGAGCGCGCGCGCGCGCGCGCGCG 1156
 Db 432 TATTCGCGATCGAGCGCGCTATAGTTTAGAGGAGCGCGCGCGCGCGCGCGCGCG 373
 QY 1157 GTGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1187
 Db 372 GTGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 342

RESULT 14
 ABQ39528/c
 ID ABQ39528 standard; DNA; 764 BP.
 XX AC ABQ39528;
 XX DT 12-JUL-2002 (first entry)
 XX OLigonucleotide for detecting cytosine methylation SEQ ID NO 26119.
 XX Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
 XX drug; side effect; cancer; central nervous system; cardiovascular;
 XX gastrointestinal; respiratory system; single nucleotide polymorphism;
 XX SNP; cell differentiation; ds.
 XX OS Homo sapiens.

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

M nucleic - nucleic search, using sw model

un on: November 28, 2002, 18:20:44 ; Search time 3019.72 Seconds
(without alignments)
11443.601 Million cell updates/sec

itle: US-09-963-333-7

effect score: 1187
sequence: 1 gatcgccactgcactcca.....gaggagcgccgctggga 1187

coring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

searched: 2054640 seqs, 14551402878 residues

total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg.hum.*
- 31: em.htg.inv.*
- 32: em.htg.other.*
- 33: em.htg.mus.*
- 34: em.htg.pln.*
- 35: em.htg.rod.*
- 36: em.htg.mam.*
- 37: em.htg.vrt.*
- 38: em.sy.*
- 39: em.htgo.hum.*
- 40: em.htgo.mus.*
- 41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1167	98.3	1186	9	HUMTS	D00517 Homo sapien
2	1167	98.3	18596	6	AR144965	AR144965 Sequence
3	1167	98.3	18596	6	AX050451	AX050451 Sequence
4	1167	98.3	18596	6	AX330682	AX330682 Sequence
5	1167	98.3	18596	6	AX330906	AX330906 Sequence
6	1167	98.3	18596	6	AX335755	AX335755 Sequence
7	1167	98.3	18596	6	AX397714	AX397714 Sequence
8	1167	98.3	18596	6	AX408943	AX408943 Sequence
9	1167	98.3	18596	6	HUMTS1	D00596 Homo sapien
10	1149.8	96.9	152711	9	AP001178	AP001178 Homo sapi
11	1006	84.8	165742	2	AC020697	AC020697 Homo sapi
12	770.6	64.9	1186	6	AX348477	AX348477 Sequence
13	710	59.8	1186	6	AX348476	AX348476 Sequence
14	299.4	25.2	1536	6	AR103889	AR103889 Sequence
15	299.4	25.2	1536	6	AX397715	AX397715 Sequence
16	299.4	25.2	1536	6	AX482020	AX482020 Sequence
17	299.4	25.2	1536	6	HSTS1	X02308 Human mRNA
18	288.8	24.3	3298	6	AX281791	AX281791 Sequence
19	278.6	23.5	1567	9	AB062290	AB062290 Homo sapi
20	278.6	23.5	1569	9	BC013919	BC013919 Homo sapi
21	275.8	23.2	1533	9	BC002567	BC002567 Homo sapi
22	228.2	19.2	233454	9	AC005703	AC005703 Homo sapi
23	225.2	19.0	80565	9	AL357374	AL357374 Human DNA
24	224.6	18.9	193487	2	AC127455	AC127455 Homo sapi
25	224.4	18.9	190871	9	AC044797	AC044797 Homo sapi
26	222.6	18.8	202827	9	AC016907	AC016907 Homo sapi
27	221.8	18.7	193487	2	AC127455	AC127455 Homo sapi
28	221.6	18.7	160536	9	AL355984	AL355984 Human DNA
29	221.2	18.6	179641	9	AC098591	AC098591 Homo sapi
30	221	18.6	150296	9	HUU91326	U91326 Homo sapien
31	220.6	18.6	97137	9	AC004745	AC004745 Homo sapi
32	220.4	18.6	132948	9	HS349A12	AL033520 Human DNA
33	219.6	18.5	146407	9	AC004953	AC004953 Homo sapi
34	219.6	18.5	195773	2	AC073134	AC073134 Homo sapi
35	219.4	18.5	104289	2	HS104D14	AL035554 Homo sapi
36	219	18.4	43497	9	HS695020A	AL049856 Human DNA
37	219	18.4	211176	2	HSDA109B7	AL121672 Homo sapi
38	218.8	18.4	156602	9	AC012487	AC012487 Homo sapi
39	218.6	18.4	148954	9	AL590640	AL590640 Human DNA
40	218.2	18.4	180346	9	AC092615	AC092615 Homo sapi
41	218	18.4	142667	9	AC125387	AC125387 Homo sapi
42	218	18.4	181610	9	AC100860	AC100860 Homo sapi
43	217.8	18.3	168006	2	AC126227	AC126227 Papio cyn
44	217.6	18.3	305000	9	HSXDPA	AL590762 Homo sapi
45	217.4	18.3	63797	2	AC127453	AC127453 Homo sapi

ALIGNMENTS

RESULT 1	LOCUS	HUMTS	1186 bp	DNA	linear	PRI 06-NOV-2001
DEFINITION	Homo sapiens	gene for thymidylate synthase, exon 1, partial cds.				
ACCESSION	D00517					
VERSION	D00517.1	GI:220133				
KEYWORDS						
SOURCE	Homo sapiens	DNA, clone:lambda HTS-1.				
ORGANISM	Homo sapiens					
REFERENCE	1 (bases 1 to 1186)					
AUTHORS	Takeishi, K., Kaneda, S., Ayusawa, D., Shimizu, K., Gotoh, O. and Seno, T.					
TITLE	Human thymidylate synthase gene: isolation of phage clones which					

cover a functionally active gene and structural analysis of the region upstream from the translation initiation codon
J. Biochem. 106 (4), 575-583 (1989)

90110051
These data kindly submitted in computer readable form by: Keiichi Takeishi
University of Shizuoka School of Food and Nutritional Sciences 395
Yada
Shizuoka-shi,
Shizuoka-ken 422
Japan
Phone: 0542-64-5540
Fax: 0542-64-5099.

FEATURES
source
1. .1186
Location/Qualifiers
/organism="Homo sapiens"
/db xref="taxon:9606"
/chromosome="18"
/clone="lambda HTS-1"
/note="Sau3AI site-from a genomic library of the mouse cell transformant CO"
115..400
repeat_unit
/note="Alu sequence"

prim_transcript
751..>1186
/notes="thymidylate synthase mRNA and intron"

exon
751..1128
/number=1
repeat_unit
785..812
/notes="inverted sequence to TTR"

repeat_unit
827..916
/notes="triple tandem repeat (TTR)"
924..>1128
/EC_number="2.1.1.45"

CDS
/note="exon part"
/codon_start=1
/product="thymidylate synthase"
/protein_id="BAA00404.1"
/db_xref="GI:2160415"

intron
/translation="MPVAGSLPRPLPPAAQERDAEPHPHGLQYLQIQLHRCG
VRKDRGTGTLGVFMQARYSLR"
1129..>1186
/number=1

BASE COUNT 224 a 386 c 347 g 229 t
ORIGIN
Query Match 98.3%; Score 1167; DB 9; Length 1186;
Best Local Similarity 99.2%; Pred. No. 4.6e-201;
Matches 1178; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

Qy 1 GATCGGCGCAGTCCAGCTCGGTGAGAGAGCGAGCTCTGTCTCAAAAAA 60
Db 1 GATCGGCGCAGTCCAGCTCGGTGAGAGAGCGAGCTCTGTCTCAAAAAA 60

Qy 61 AAAAAAGCGCGAGGCTCAACAAAAAAGCTCGGAAGCCCTGGGAGCTTTT 120
Db 61 AAAAAAGCGCGAGGCTCAACAAAAAAGCTCGGAAGCCCTGGGAGCTTTT 120

Qy 121 TTTTCTTTTCTTTTCTTTGAGCAGTCTCTGTGCTCCCGAGCTGGAGTAC 180
Db 121 TTTTCTTTTCTTTTCTTTGAGCAGTCTCTGTGCTCCCGAGCTGGAGTAC 180

Qy 181 CGGATCTTGCTCACTCAACCTCTGCTCCCGAGCTTCAAGCAATCTTCT 240
Db 181 CGGATCTTGCTCACTCAACCTCTGCTCCCGAGCTTCAAGCAATCTTCT 240

Qy 241 TCCAGTAGCCACACCGCCAGTAAATTTTGTANTTTTGTAGAGACGGGTT 300
Db 241 TCCAGTAGCCACACCGCCAGTAAATTTTGTANTTTTGTAGAGACGGGTT 300

Qy 301 CATGTTGCTCCAGGCTGTCTGAACTCTGACTCAGGTGATCCACCGCTCG 360
Db 301 CATGTTGCTCCAGGCTGTCTGAACTCTGACTCAGGTGATCCACCGCTCG 360

Qy 361 CAAAGTACTAGGATTACAGGCGTGAGCCACCGCGTCCAGCGCTTGTATCA 420
Db 361 CAAAGTACTAGGATTACAGGCGTGAGCCACCGCGTCCAGCGCTTGTATCA 420
Qy 421 AGTAGAAAAGCTGCAATATATACCACTTCTGCTTCNGTTGCTNTTCAGTGAAGAAATGG 480
Db 421 AGTAGAAAAGCTGCAATATATACCACTTCTGCTTCNGTTGCTTCAGTGAAGAAATGG 479
Qy 481 AAATGCAATNCNTTATTTAGTTTGTAGAAAACAGATCTCAACACAGAGTTTGTGNGACAAG 540
Db 480 AAATGCAATNCNTTATTTAGTTTGTAGAAAACAGATCTCAACACAGAGTTTGTGNGACAAG 539
Qy 541 ACCGAGGAAAACGTGGGAACTGTGTCTGTCTGTCTGTAGAGAGGCGCGGTCCAGCAGCGG 600
Db 540 ACCGAGGAAAACGTGGGAACTGTGTCTGTCTGTCTGTAGAGAGGCGCGGTCCAGCAGCGG 599
Qy 601 TTCCCAAGAGGCGAGTCTCTTCCCGCCACCGACCTGCNTCCAGGTTCCCGGTTTCT 660
Db 600 TTCCCAAGAGGCGAGTCTCTTCCCGCCACCGACCTGCNTCCAGGTTCCCGGTTTCT 659
Qy 661 AAGACTCTCAGTGTGCGCTTGGGCTCCGTTCTGTGCGCACACCGCTGCTCTGCGTTTC 720
Db 660 AAGACTCTCAGTGTGCGCTTGGGCTCCGTTCTGTGCGCACACCGCTGCTCTGCGTTTC 719
Qy 721 CCCTCGCGCACCGCTCTCTAGAGCGGCGCGCGCGCGCGCGCGCGCGCGAGAGAGGCG 780
Db 720 CCCTCGCGCACCGCTCTCTAGAGCGGCGCGCGCGCGCGCGCGCGCGCGAGAGAGGCG 779
Qy 781 GAGCGCGGAGCGCGCGCGGAAAAGCGCGCGGAGGCGTCTGCGCACCGCGCGCACTTGG 840
Db 780 GAGCGCGGAGCGCGCGCGGAAAAGCGCGGAGGCGTCTGCGCACCGCGCGCACTTGG 839
Qy 841 CTGCGCTCCGCTCCG 900
Db 840 CTGCGCTCCGCTCCG 899
Qy 901 CTGCGCTCCG 960
Db 900 CTGCGCTCCG 959
Qy 961 TTGCG 1020
Db 960 TTGCG 1019
Qy 1021 TACTGCGCGCAGATCCAAACATCTCTCGCTGCGCGCGTTCAGGAGGAGCGACCGCGCGCG 1080
Db 1020 TACTGCGCGCAGATCCAAACATCTCTCGCTGCGCGCGTTCAGGAGGAGCGACCGCGCGCG 1079
Qy 1081 ACCGCGACCTCTCGGTATTCGCGATGCGAGCGCGCTACAGCTTGAAGGTGACCGCGCG 1140
Db 1080 ACCGCGACCTCTCGGTATTCGCGATGCGAGCGCGCTACAGCTTGAAGGTGACCGCGCGCG 1139
Qy 1141 GGCGCTGCGGAGCGGTTGGGCGGAGAGGAGGAGGCGCGCGCGCGCGCGCGCGCGCG 1187
Db 1140 GGCGCTGCGGAGCGGTTGGGCGGAGAGGAGGAGGCGCGCGCGCGCGCGCGCGCGCG 1186

RESULT 2
AR144965 18596 bp DNA linear PAT 09-AUG-2001
LOCUS
DEFINITION
Sequence 11 from patent US 6210950.
ACCESSION
AR144965
VERSION
AR144965.1 GI:15106832
KEYWORDS
SOURCE
Unknown.
ORGANISM
Unclassified.
REFERENCE
1 (bases 1 to 18596)
Johnson, W.G. and Stenroos, B.Scott.
METHODS
Methods for diagnosing, preventing, and treating developmental disorders due to a combination of genetic and environmental factors
PATENT
Patent: US 6210950-A 11 03-APR-2001;
JOURNAL
Location/Qualifiers
FEATURES

```
source 1. 18596
/organism="unknown"
BASE COUNT 4521 a 3991 c 4479 g 5605 t
ORIGIN
Query Match 98.3%; Score 1167; DB 6; Length 18596;
Best Local Similarity 99.2%; Pred. No. 3.8e-201;
Matches 1178; Conservative 0; Mismatches 8; Indels 1; Gaps 1;
y 1 GATCGCGCCTCAGTCCAGCTCCAGCTGGGTGAGAGACGAGACTCTGTCTCAAAAAA 60
b 78 GATCGCGCCTCAGTCCAGCTCCAGCTGGGTGAGAGACGAGACTCTGTCTCAAAAAA 137
y 61 AAAAAAGACCGCAGGCTCAAAACAAAAACCTCGAAAGACCTCGGAGGCTCTTTT 120
b 138 AAAAAAGACCGCAGGCTCAAAACAAAAACCTCGAAAGACCTCGGAGGCTCTTTT 197
y 121 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 180
b 198 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 257
y 181 CGGATCTGGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTG 240
b 258 CGGATCTGGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTG 317
y 241 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 300
b 318 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 377
y 301 CATGTGTCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
b 378 CATGTGTCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 437
y 361 CAAAGTACTAGGATTAAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 420
b 438 CAAAGTACTAGGATTAAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 497
y 421 AGTAGAAAAGCTGCAATTATACCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCT 480
b 498 AGTAGAAAAGCTGCAATTATACCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCT 556
y 481 AATGCAATCMTTATTTAGTTTGTAGAAAACAGATCTCAAAACAGAGTTTGTGACAAG 540
b 557 AATGCAATCMTTATTTAGTTTGTAGAAAACAGATCTCAAAACAGAGTTTGTGACAAG 616
y 541 ACCGAGGAAAAGCTGGGAACTGTGCTGCTGGCTGTAGAGAGCGCGCTCGACACAGG 600
b 617 ACCGAGGAAAAGCTGGGAACTGTGCTGCTGGCTGTAGAGAGCGCGCTCGACACAGG 676
y 601 TTCCAAAGGCGCAGTCTCTCCGCGCAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
b 677 TTCCAAAGGCGCAGTCTCTCCGCGCAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCT 736
y 661 AAGACTCTCAGCTGTGGCTCTGGCTCTGGCTCTGGCTCTGGCTCTGGCTCTGGCTCT 720
b 737 AAGACTCTCAGCTGTGGCTCTGGCTCTGGCTCTGGCTCTGGCTCTGGCTCTGGCTCT 796
y 721 CCCCTGGCGCAGCTCTCTAGAGCGGGGCGCGCGAGACCCCGCGAGCAGAGAGGCG 780
b 797 CCCCTGGCGCAGCTCTCTAGAGCGGGGCGCGCGAGACCCCGCGAGCAGAGAGGCG 856
y 781 GAGCGCGGAGACCGCGCGGAAAAGCGCGCGGAGAGGCTGCTGCTGCTGCTGCTGCTG 840
b 857 GAGCGCGGAGACCGCGCGGAAAAGCGCGCGGAGAGGCTGCTGCTGCTGCTGCTGCTG 916
y 841 CTTGCTCTCGCTCGCGCGCGCACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
b 917 CTTGCTCTCGCTCGCGCGCGCACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 976
y 901 CTTGCTCTCGCTCGCGCGCGCACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
b 977 CTTGCTCTCGCTCGCGCGCGCACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1036
```


1 961 TTGCCCCCGCCGACAGAGCGGAGCGCCGAGCGCGCTCCGCGCAGCGGAGCTGCAG 1020
1037 TTGCCCCCGCCGACAGAGCGGAGCGCCGAGCGCGCTCCGCGCAGCGGAGCTGCAG 1096
1021 TACCTGGGCGAGATCCAAACACATCTCTCCGCTGCGCGCTCAGGAAGCAGCAGCGGC 1080
1097 TACCTGGGCGAGATCCAAACACATCTCTCCGCTGCGCGCTCAGGAAGCAGCAGCGGC 1156
1081 ACCGCAACCTGTCGCTATTGCGGATGTCAGGCGCGCTTACAGCTGAGAGGTGACGCGCG 1140
1157 ACCGCAACCTGTCGCTATTGCGGATGTCAGGCGCGCTTACAGCTGAGAGGTGACGCGCG 1216
1141 GCGCCCTGCGGAGCGGCTGCGGAGGAGGAGCGCGCTGGGGA 1187
1217 GCGCCCTGCGGAGCGGCTGCGGAGGAGGAGCGCGCTGGGGA 1263

RESULT 5
X330906 AX330906 18596 bp DNA linear PAT 09-JAN-2002
OCUS Sequence 1415 from Patent WO0194629.
CESSION AX330906
BESION AX330906.1 GI:18103885
EYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini, Homnidae; Homo.

REFERENCE
1
AUTHORS Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
Horrigan, S., Soppet, D.R. and Weaver, Z.
TITLE Cancer gene determination and therapeutic screening using signature
gene sets
JOURNAL Patent: WO 0194629-A 1415 13-DEC-2001;
Avalon Pharmaceuticals (US)
FEATURES Location/Qualifiers
source
1..18596
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 4521 a 3991 c 4479 g 5605 t
ORIGIN

Query Match 98.3%; Score 1167; DB 6; Length 18596;
Best Local Similarity 99.2%; Pred. No. 3.8e-201;
Matches 1178; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

2Y 1 GATCGCGCACTGCACTCCAGCTCCAGCTGGGTGAGAGCGAGACTCTGTCTCAAAAAA 60
Db 78 GATCGCGCACTGCACTCCAGCTGGGTGAGAGCGAGACTCTGTCTCAAAAAA 137
QY 61 AAAAAGACCGCAGGCTCAAAACAAAAAAGCTCGGAAAGCTCGGAAAGCTCGGAAAGCT 120
Db 138 AAAAAGACCGCAGGCTCAAAACAAAAAAGCTCGGAAAGCTCGGAAAGCTCGGAAAGCT 197
QY 121 TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT 180
Db 198 TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT 257
QY 181 CGATCTTGGCTCACTGGAACCTTCTGCTCCAGCTTCAAGCAATCTTCTGCTCAGCC 240
Db 258 CGATCTTGGCTCACTGGAACCTTCTGCTCCAGCTTCAAGCAATCTTCTGCTCAGCC 317
QY 241 TCCCAAGTAGCCACCGCCAGCTAAATTTTGTANTTTTAGTAGAGCGGGGTTTCAC 300
Db 318 TCCCAAGTAGCCACCGCCAGCTAAATTTTGTANTTTTAGTAGAGCGGGGTTTCAC 377
QY 301 CATGTTGTCAGGCTGCTGTAACCTCTGCACTCAGCTGATCCACCGCTCGGCCCC 360
Db 378 CATGTTGTCAGGCTGCTGTAACCTCTGCACTCAGCTGATCCACCGCTCGGCCCC 437
QY 361 CAAAGTACTAGATTAAGGCTGAGCCACCGCTCCAGCGCTCGCGGTTTAAATCA 420
Db 438 CAAAGTACTAGATTAAGGCTGAGCCACCGCTCCAGCGCTCGCGGTTTAAATCA 497

QY 421 AGTAGAAGAGCTGCATTTATACCACTTCTTCNTTTCNTTTCAGTGAGAACGAAGAATCG 480
Db 498 AGTAGAAGAGCTGCATTTATACCACTTCTTCNTTTCNTTTCAGTGAGAACGAAGAATCG 556
QY 481 AAATGCAATTCNTTATTTAGTTGTAGAAACAGATCTCAAAACAGCAGTCTTTGTGACAAG 540
Db 557 AAATGCAATTCNTTATTTAGTTGTAGAAACAGATCTCAAAACAGCAGTCTTTGTGACAAG 616
QY 541 ACCCGAGGAAAACGTGGGAAATGTGTGCTGTAGAGAAAGGCGCGGTGCAACAGCGG 600
Db 617 ACCCGAGGAAAACGTGGGAAATGTGTGCTGTAGAGAAAGGCGCGGTGCAACAGCGG 676
QY 601 TTCCCAAGAGGCGCAGTCTCTCCGCGCACCGCAGCTGNTCCAGGTTCCCGGGTTTCCT 660
Db 677 TTCCCAAGAGGCGCAGTCTCTCCGCGCACCGCAGCTGNTCCAGGTTCCCGGGTTTCCT 736
QY 661 AAGACTCTCAGCTGTGGCTCCGCTTCTGTGCGCACACCCGCTGGCTCTCTGCGTTTC 720
Db 737 AAGACTCTCAGCTGTGGCTCCGCTTCTGTGCGCACACCCGCTGGCTCTCTGCGTTTC 796
QY 721 CCCCTGGCGCAGCTCTCTAGAGCGGGGCGCGCGCAACCCCGCGCAGCAGGAAGAGCG 780
Db 797 CCCCTGGCGCAGCTCTCTAGAGCGGGGCGCGCGCAACCCCGCGCAGCAGGAAGAGCG 856
QY 781 GAGCGCGGAGCGCGCGCGGAAAAGGCGCGCGAAGGGGTCTTGCACACCGCGCACTTGG 840
Db 857 GAGCGCGGAGCGCGCGCGGAAAAGGCGCGCGAAGGGGTCTTGCACACCGCGCACTTGG 916
QY 841 CCTGCTCCGCTCCGCGCGCACCTTGGCTGCTCCGTCGCGCGCGCGCTTGCCTTG 900
Db 917 CCTGCTCCGCTCCGCGCGCACCTTGGCTGCTCCGTCGCGCGCGCGCTTGCCTTG 976
QY 901 CCTGCTCCGCTCCGCGCGCACCTTGGCTGCTCCGTCGCGCGCGCGCTTGCCTTG 960
Db 977 CCTGCTCCGCTCCGCGCGCACCTTGGCTGCTCCGTCGCGCGCGCGCTTGCCTTG 1036
QY 961 TTGCCCCCGCCGACAGAGCGGAGCGCGAGCTCGCTCCGTCGCGCGCGCGCTTGCCTTG 1020
Db 1037 TTGCCCCCGCCGACAGAGCGGAGCGCGAGCTCGCTCCGTCGCGCGCGCGCTTGCCTTG 1096
QY 1021 TACCTGGGCGAGATCCAAACACATCTCTCCGCTGCGCGCTCAGGAAGCAGCAGCGGC 1080
Db 1097 TACCTGGGCGAGATCCAAACACATCTCTCCGCTGCGCGCTCAGGAAGCAGCAGCGGC 1156
QY 1081 ACCGCAACCTGTCGCTATTGCGGATGTCAGGCGCGCTTACAGCTGAGAGGTGACGCGCG 1140
Db 1157 ACCGCAACCTGTCGCTATTGCGGATGTCAGGCGCGCTTACAGCTGAGAGGTGACGCGCG 1216
QY 1141 GCGCCCTGCGGAGCGGCTGCGGAGGAGGAGCGCGCTGGGGA 1187
Db 1217 GCGCCCTGCGGAGCGGCTGCGGAGGAGGAGGAGCGCGCTGGGGA 1263

RESULT 6
AX335755 18596 bp DNA linear PAT 09-JAN-2002
LOCUS Sequence 6264 from Patent WO0194629.
DEFINITION AX335755
ACCESSION AX335755
VERSION AX335755.1 GI:18126474
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1
AUTHORS Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
Horrigan, S., Soppet, D.R. and Weaver, Z.
TITLE Cancer gene determination and therapeutic screening using signature
gene sets
JOURNAL Patent: WO 0194629-A 6264 13-DEC-2001;
Avalon Pharmaceuticals (US)
FEATURES Location/Qualifiers

```
source 1..18596
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 4521 a 3991 c 4479 g 5605 t
ORIGIN
Query Match 98.3%; Score 1167; DB 6; Length 18596;
Best Local Similarity 99.2%; Pred. No. 3.8e-201;
Matches 1178; Conservative 0; Mismatches 8; Indels 1; Gaps 1;
QY 1 GATCGCGCACTGCACTCCAGCTGGGTGAGAGCGAGACTCTGTCTCAAAAAA 60
Db 78 GATCGCGCACTGCACTCCAGCTGGGTGAGAGCGAGACTCTGTCTCAAAAAA 137
QY 61 AAAAAAGACCGCGGCTCAAAACAAAACCTCGGAAAGCCCTGCGGCTCTTTT 120
Db 138 AAAAAAGACCGCGGCTCAAAACAAAACCTCGGAAAGCCCTGCGGCTCTTTT 197
QY 121 TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT 180
Db 737 AAGACTCTAGCTGTGGGCTGGGCTCCGTTCTGTGCGCACACCCGTGGCTTC 796
QY 721 CCGCTGCGCAGCTCTTAGAGCGGGGCGCGGAGCCCGCGAGAGAGAGAGGCG 780
Db 797 CCGCTGCGCAGCTCTTAGAGCGGGGCGCGGAGCCCGCGAGAGAGAGAGGCG 856
QY 781 GAGCGGAGAGCGCGCGGAGAGAGCGCGGAGAGAGAGAGAGAGAGAGAGAG 840
Db 857 GAGCGGAGAGCGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 916
QY 841 CCGCTGCGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTG 900
Db 917 CCGCTGCGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTG 976
QY 901 CCGCTGCGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTG 960
Db 977 CCGCTGCGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTG 1036
```

```
QY 961 TTGCCCCCGCGCGCACAGGAGCGGCGAGCGCGCGCTCCGCGCACAGGAGCTGCAG 1020
Db 1037 TTGCCCCCGCGCGCACAGGAGCGGCGAGCGCGCGCTCCGCGCACAGGAGCTGCAG 1096
QY 1021 TACCTGGGCGAGATCCAAACATCTCTCCGCTGCGGCGCTCAGGAGAGACGACCGACGGGC 1080
Db 1097 TACCTGGGCGAGATCCAAACATCTCTCCGCTGCGGCGCTCAGGAGAGACGACCGACGGGC 1156
QY 1081 ACGGCAACCTCTGCTGGTATTCGCGATGACGGCGCTACGCTTGAGAGGTGACCGCGG 1140
Db 1157 ACGGCAACCTCTGCTGGTATTCGCGATGACGGCGCTACGCTTGAGAGGTGACCGCGG 1216
QY 1141 GCGCCCTCGGAGACGGGTGCGGAAAGAGGAGGCGCGGCTGGGGA 1187
Db 1217 GCGCCCTCGGAGACGGGTGCGGAAAGAGGAGGCGCGGCTGGGGA 1263
RESULT 7
AX397714 18596 bp DNA linear PAT 18-MAY-2002
LOCUS
DEFINITION Sequence 140 from Patent WO210366.
ACCESSION AX397714
VERSION AX397714.1 GI:21068436
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1
AUTHORS Chen, H., Freilmer, N. B. and Novak, T.
TITLE Methods and compositions for diagnosing and treating chromosome-18P
related disorders
JOURNAL Patent: WO 0210366-A 140 07-FEB-2002;
Millennium Pharmaceuticals, Inc. (US) ; The Regents of The
University of California (US)
FEATURES
source 1..18596
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 4521 a 3991 c 4479 g 5605 t
ORIGIN
Query Match 98.3%; Score 1167; DB 6; Length 18596;
Best Local Similarity 99.2%; Pred. No. 3.8e-201;
Matches 1178; Conservative 0; Mismatches 8; Indels 1; Gaps 1;
QY 1 GATCGCGCACTGCACTCCAGCTGGGTGAGAGCGAGACTCTGTCTCAAAAAA 60
Db 78 GATCGCGCACTGCACTCCAGCTGGGTGAGAGCGAGACTCTGTCTCAAAAAA 137
QY 61 AAAAAAGACCGCGGCTCAAAACAAAACCTCGGAAAGCCCTGCGGCTCTTTT 120
Db 138 AAAAAAGACCGCGGCTCAAAACAAAACCTCGGAAAGCCCTGCGGCTCTTTT 197
QY 121 TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT 180
Db 737 AAGACTCTAGCTGTGGGCTGGGCTCCGTTCTGTGCGCACACCCGTGGCTTC 796
QY 721 CCGCTGCGCAGCTCTTAGAGCGGGGCGCGGAGCCCGCGAGAGAGAGAGGCG 780
Db 797 CCGCTGCGCAGCTCTTAGAGCGGGGCGCGGAGCCCGCGAGAGAGAGAGGCG 856
QY 781 GAGCGGAGAGCGCGGAGAGAGCGCGGAGAGAGAGAGAGAGAGAGAGAG 840
Db 857 GAGCGGAGAGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 916
QY 841 CCGCTGCGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTG 900
Db 917 CCGCTGCGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTG 976
QY 901 CCGCTGCGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTG 960
Db 977 CCGCTGCGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTG 1036
```


QY 961 TTGCCCCCGCGGCACAGGAGCGGACGCCGAGCGCGTCCGCCGACCGGAGCTGCAG 1020
 |||||
 Db 1037 TTGCCCCCGCGGCACAGGAGCGGACGCCGAGCGCGTCCGCCGACCGGAGCTGCAG 1096
 |||||
 QY 1021 TACCTGGGCGGATCACAACATCTCCGCTGCGGCGTCTAGGAAGACCGACCGGCGGC 1080
 |||||
 Db 1097 TACCTGGGCGGATCACAACATCTCCGCTGCGGCGTCTAGGAAGACCGACCGGCGGC 1156
 |||||
 QY 1081 ACCGCGACCTGTCTGGTATTCGCGATGCGGCGCGCTACAGCTGAGAGTGCACCGCGC 1140
 |||||
 Db 1157 ACCGCGACCTGTCTGGTATTCGCGATGCGGCGCGCTACAGCTGAGAGTGCACCGCGC 1216
 |||||
 QY 1141 GCGCCCTCGGAGCGGTGCGCGGAGGAGCGCGCGGCTGGGA 1187
 |||||
 Db 1217 GCGCCCTCGGAGCGGTGCGCGGAGGAGCGCGCGGCTGGGA 1263
 |||||

RESULT 9

HUMTS1 18596 bp DNA linear PRI 14-APR-2000
 LOCUS
 DEFINITION Homo sapiens gene for thymidylate synthase, exons 1, 2, 3, 4, 5, 6,
 7, complete cds.

ACCESSION D00596

VERSION D00596.1 GI:220135

KEYWORDS thymidylate synthase.

SOURCE Homo sapiens DNA, clone: lambdaHTS-1 and lambdaHTS-3.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 18596)

AUTHORS Kaneda, S., Nalbantoglu, J., Takeishi, K., Shimizu, K., Gotoh, O.,

Seno, T. and Aoyama, D.

TITLE Structural and functional analysis of the human thymidylate

synthase gene

J. Biol. Chem. 265 (33), 20277-20284 (1990)

MEDLINE 91056070

COMMENT These data kindly submitted in computer readable form by: Sumiko

Kaneda

National Institute of Genetics

1111 Yata

Mishima 411

Japan

Phone: +81-559-72-2732

Fax: +81-559-71-3651.

Location/Qualifiers

1. 18596

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="18"

/map="18p11.32"

/clones="lambdaHTS-1 and lambdaHTS-3"

repeat_unit 1. .148

/note="Alu sequence"

repeat_unit 202. .477

/note="Alu sequence"

prim_transcript 822. .16246

/note="thymidylate synthase mRNA and introns"

prim_transcript 824. .16246

/note="thymidylate synthase mRNA and introns"

prim_transcript 828. .16246

/note="thymidylate synthase mRNA and introns"

prim_transcript 841. .16246

/note="thymidylate synthase mRNA and introns"

repeat_unit 862. .889

/note="inverted repeat"

repeat_unit 904. .993

/note="triple tandem repeats"

join(1001..1205,2895..2968,5396..5570,11843..11944,

13449..13624,14133..14204,15613..15750)

/EC number="2.1.1.45"

/codon_start=1

/product="thymidylate synthase"

/protein_id="BAA00472.1"

/db_xref="GI:220136"

/translation="MPVAGSELPRRLPPAAQERDAERPPHGELOVIGIQLHRCG

VRDDTGTGTLVFGMQARYSLRDBFPLITTKRVFKGLLELLNFKGSTNAKELS

SKGKTIWDANGSRDFLDSLGSTREGLGPVGFQWRHFGAERYDMESYSGQVDQ

LQRVITIKTPDDERLIIMCAWNPRLPLMALPPCHALCQFVNVNSELSCQLXQSGD

MGVGFNFNIATYALLTYIAHITGLKPGDFIHTLGAHLYLNHIEPLKIQLQREPRFF

PKLRIILRKVEKIDDFKAFDFQIEGYNPHPTIKEMAV"

<1001..1205

/number=1

1206..2894

/number=1

1660..1665

/bound_moiety="Spl"

1890..1895

/bound_moiety="Spl"

1895..1900

/bound_moiety="Spl"

2895..2968

/number=2

2969..5395

/number=2

3049..3347

/note="Alu sequence"

4953..5245

/note="Alu sequence"

5396..5570

/number=3

5571..11842

/number=3

5768..9405

/note="LI sequence"

10125..10389

/note="reiterated motif of GGA/TCGAT"

11843..11944

/number=4

11945..13448

/number=4

12135..12397

/note="Alu sequence"

12567..12743

/note="Alu sequence"

12807..13084

/note="Alu sequence"

13449..13624

/number=5

13625..14132

/number=5

13875..14010

/note="Alu sequence"

14133..14204

/number=6

14205..15612

/number=6

14531..14832

/note="Alu sequence"

15613..15750

/number=7

15934..15939

polyA signal

16228..16233

polyA signal

17262..17553

repeat_unit

/note="Alu sequence"

BASE COUNT 4521 a 3991 c 4479 g 5605 t

ORIGIN

Query Match 98.3%; Score 1167; DB 9; Length 18596;

Best Local Similarity 99.2%; Pred. No. 3.8e-201;

Matches 1178; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

QY 1 GATCGCGCCTGCACCTCCAGCTGGTGGTGAGAGCGAGACTCTGTCTCAAAAAAAA 60

Db 78 GATCGCGCCTGCACCTCCAGCTGGTGGTGAGAGCGAGACTCTGTCTCAAAAAAAA 137

61	AAAAGACCGCCAGGGCTCAAAACAAAAACCTCGAAAAAGCCCTGGCGGTCTTTTTTTTT	120
y		
138	AAAAGACCGCCAGGGCTCAAAACAAAAACCTCGAAAAAGCCCTGGCGGTCTTTTTTTTT	137
b		
121	TTTTTTTTTTTTTTTTTTTTTTGGGACAGCTCTTGCTCTGTGCGCCAGGCTGGAGTGACAAATGSET	180
y		
198	TTTTTTTTTTTTTTTTTTTTTTGGGACAGCTCTTGCTCTGTGCGCCAGGCTGGAGTGACAAATGSET	257
b		
181	CGGATCTTTGGCTCACTGCAACCTCTTGCCCTCCACGGTTCAAGCAATTTCTTCTGCTCTCAGCC	240
y		
258	CGGATCTTTGGCTCACTGCAACCTCTTGCCCTCCACGGTTCAAGCAATTTCTTCTGCTCTCAGCC	317
b		
241	TCCCAAGTAGCCACCGCCGACGCTAAATTTTTGTANTTTTTAGTAGAGACGGGGTTTTCAC	300
y		
318	TCCCAAGTAGCCACCGCCGACGCTAAATTTTTGTANTTTTTAGTAGAGACGGGGTTTTCAC	377
b		
301	CATGTTGTCACAGGCTGGCTCTGAACTCTCTGACCTCAGGTGATCCACCCGCTCTGCGCCCTC	360
y		
378	CATGTTGTCACAGGCTGGCTCTGAACTCTCTGACCTCAGGTGATCCACCCGCTCTGCGCCCTC	437
b		
361	CAAAAGTACTAGGATTACAGCGGTGAGCCACCGCTCTCAGGCCCTGGCGGTTTTTAATCA	420
y		
438	CAAAAGTACTAGGATTACAGCGGTGAGCCACCGCTCTCAGGCCCTGGCGGTTTTTAATCA	497
b		
421	AGTAGAAAAGCTGCATATATACCTCTTGCTCTGTTGCTTTTCACTGAGAGACGAGAAATGG	480
y		
498	AGTAGAAAAGCTGCATATATACCTCTTGCTCTGTTGCTTTTCACTGAGAGACGAGAAATGG	556
b		
481	AAATGCAAAATCNCCTTATTAGTCTGTAGAAACAGATCTCAAACACAGTCTTTGTGACAAAG	540
y		
557	AAATGCAAAATCNCCTTATTAGTCTGTAGAAACAGATCTCAAACACAGTCTTTGTGACAAAG	616
b		
541	ACCGCAGAAAACGTGGGAACTGTGCTGTGCTGTAGAGAAAGCGCGCTCGACCAAGG	600
y		
617	ACCGCAGAAAACGTGGGAACTGTGCTGTGCTGTAGAGAAAGCGCGCTCGACCAAGG	676
b		
601	TTCCCAAGAGGGCGAGTCCCTTCCGCGCCACCGCACCTGCTCAGGTTCCCGGGTTCCT	660
y		
677	TTCCCAAGAGGGCGAGTCCCTTCCGCGCCACCGCACCTGCTCAGGTTCCCGGGTTCCT	736
b		
661	AAGACTCTCAGCTGTGGCCCTTGGGCTCCGTTTGTGACCAACCCGTGGCTCTCTGCGTTTC	720
y		
737	AAGACTCTCAGCTGTGGCCCTTGGGCTCCGTTTGTGACCAACCCGTGGCTCTCTGCGTTTC	796
b		
721	CCCTGCGGCAACGCTCTCTAGAGCGGGGGCCCGCGACCCCGCGAGCAGAGAGAGGCG	780
y		
797	CCCTGCGGCAACGCTCTCTAGAGCGGGGGCCCGCGACCCCGCGAGCAGAGAGAGGCG	856
b		
781	GAGCGGGGAGCGCCCGCGGAAAAGGCGCGGAAAGGGTCTCTGCCACCGCGGCACCTTGG	840
y		
857	GAGCGGGGAGCGCCCGCGGAAAAGGCGCGGAAAGGGTCTCTGCCACCGCGGCACCTTGG	916
b		
841	CTTGCTCTCGTCCGCGCGGCACATCTGAGCTGCTCCGTCGCGCGCGCCACTTCGCGCTG	900
y		
917	CTTGCTCTCGTCCGCGCGGCACATCTGAGCTGCTCCGTCGCGCGCGCCACTTCGCGCTG	976
b		
901	CCTCGGTCGCCCGCCCGCGCGCCATGCTGTGCGCGGCTCGGAGCTCGCGCGCGGCC	960
y		
977	CCTCGGTCGCCCGCCCGCGCGCCATGCTGTGCGCGGCTCGGAGCTCGCGCGCGGCC	1036
b		
961	TTGCGCCCGCCCGCCACAGGAGCGGGAGCGCGAGCGCGGTCCGCGGACGCGGAGCTCGAG	1020
y		
1037	TTGCGCCCGCCCGCCACAGGAGCGGGAGCGCGAGCGCGGTCCGCGGACGCGGAGCTCGAG	1096
b		
1021	TACCTGGGGCAGATCCAAACATCTCCGCTGCGGCGCTCAGAGAGAGACGACGACGCGGC	1080
y		
1097	TACCTGGGGCAGATCCAAACATCTCCGCTGCGGCGCTCAGAGAGAGACGACGACGCGGC	1156
b		
1081	ACCGGCAACCTGTCCGTTATTCGGGCACTCAGGCGCGCTTACAGCTGAGAGGTTGACGCGCG	1140
y		
1157	ACCGGCAACCTGTCCGTTATTCGGGCACTCAGGCGCGCTTACAGCTGAGAGGTTGACGCGCG	1216
b		
1141	GGCCCTCTGGGACGGGTGGCGGGAAGAGAGGAGCGCGCTGGGA	1187
y		

Df 1217 GSCCCCTGGGACGGTGGCGGAGGAGGAGCGCGCTGGGA 1263
|||||

RESULT 10 AP001178 152711 bp DNA linear PRI 11-JUL-2002
AP001178 Homo sapiens genomic DNA, chromosome 18 clone:RP11-806L2, complete
LOCUS sequence.
DEFINITION AP001178.4 GI:21728142
HTG.
ACCESSION Homo sapiens DNA, clone:RP11-806L2.
VERSION Homo sapiens
KEYWORDS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
SOURCE Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE Homo sapiens genomic DNA
JOURNAL Published Only in Database (2000)
REFERENCE 2 (bases 1 to 152711)
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
DIRECT SUBMISSION
TITLE Direct Submission
JOURNAL Submitted (14-FEB-2000) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC); Japan
1-7-22 Suehiro-chou,Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT On Jul 10, 2002 this sequence version replaced gi:9229948.
FEATURES
Location/Qualifiers
1..152711
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="18"
/map="18p"
/clone="RP11-806L2"

BASE COUNT 42612 a 32833 c 33397 g 43869 t
ORIGIN

Query Match 96.9%; Score 1149.8; DB 9; Length 152711;
Best Local Similarity 98.9%; Pred. No. 4.3e-198;
Matches 1174; Conservative 0; Mismatches 10; Indels 3; Gaps 2

QY 1 GATCGGCCACTGCACCTCCAGCTGGGTGAGAGCGAGACTCTGTCAAAAAAAAAA 60
Db 84508 GATCGCACCATTCGACTCCAGCTGGGTGAGAGCGAGACTCTGTCTCAAAAAAAAAA 84567

QY 61 AAAAAAGCCGACGAGGGCTCAAACAATAAACCTCGGAAAGCCCTGGGGTCTTTTTTTT 120
Db 84568 AAAAAGACCGCAGGGCTCAAACAATAAACCTCGGAAAGCCCTGGGGTCT--TTTTTTT 84625

QY 121 TTTTTTTTTTTTTTTTTTTTGGAAGAGCTTGTCTGTGCGCCAGGCTGGAGTACAATGGT 180
Db 84626 TTTTTTTTTTTTTTTTTTTTGGAAGAGCTTGTCTGTGCGCCAGGCTGGAGTACAATGGT 84685

QY 181 CGGATCTTGCTCACTGCAACCTCTGCTCCAGGTTCAAGCAATTCTTCGCCTCAGCC 240
Db 84686 CGGATCTTGCTCACTGCAACCTCTGCTCCAGGTTCAAGCAATTCTTCGCCTCAGCC 84745

QY 241 TCCCAGTAGCCACCACGCCAGCTAATTTTGTAMTTTGTAGTAGACGGGGTTTCAC 300
Db 84746 TCCCAGTAGCCACCACGCCAGCTAATTTTGTACTTTTGTAGTAGACGGGGTTTCAC 84805

QY 301 CATGTTGTCCAGGCTGCTNGAACTCCTGACCTCAGGTGATCCACCCGCTCGGCCCC 360
Db 84806 CATGTTGTCCAGGCTGCTTGAATCTCTGACCTCAGGTGATCCACCCGCTCGGCCCC 84865

QY 361 CAAAGTACTAGGATTAACGGCGTGAGCCAACCGGCTCCAGGCGCTTGGCGGTTTAAATCA 420
Db 84866 CAAAGTACTAGGATTAACGGCGTGAGCCAACCGGCTCCAGGCGCTTGGCGGTTTAAATCA 84925

Q7 1141 GGCCCTTGGCGGGAAGGGTGGCGGGAAGGAGGCGCGGCTGGGGA 1187


```
QY 182 GGATTTGGCTACTGCAACCTCTGCTCCAGGTTCAAGCAATCTTCTGCTCAGCCT 241
Db 1095 GAACTTTAACTACTACCACTCTACCTCCCAAAATTCAAACAATCTTCTACCTCAACCT 946
QY 242 CCCAAGTAGCCACACGCGCCAGCTAAATTTTGTANTTTTAGTAGAGAGGGGTTTCAAC 301
Db 945 CCCAAATAACCAACGCGCCAACTAAATTTTATATCTTTTAAATAAAGAAATTTCAAC 886
QY 302 ATGTTTCCAGGCTGTCTINGAATCTCTGACCTCAGGTGATCCACCGGCTCGGCCCCC 361
Db 885 ATATTATCCAACTAATCTTAACTCTTAACCTCAAAATATCCACCGGCTCGGCCCCC 826
QY 362 AAGTACTAGGATTAAGCGGTGAGCCACCGCTCCAGCGCTCGCGGCTTTTAAATCAA 421
Db 825 AAAATATCTAAATTAAGAAACCAACCGCTCCAGCGCTCAAGCTTAAATCAA 766
QY 422 GTAGAAAGCTGCAATATACCACTTGTCTTCGTTGCTTCAAGTGAAGAAATGGA 481
Db 765 ATAAATAAATCAATATATACCACTTACTTCGATTAC-TTCAATAAATAAGAAATATAA 707
QY 482 AATGCAATCMTCTTATTTAGTTTGTAGAAACAGATCTCAAAACAGAGTTTGTNGACAGA 541
Db 706 AATACAAATCCCTTATTAATATAAATAAAGAAATCTCAACAATAATTTTAAACAAA 647
QY 542 CGCAGGAAACCGTGGAACTGTGTCTGGCTGTAGAAAGGCGCGGTGACACACAGGT 601
Db 646 CGCAGAAACCGTGGAACTTACTACTTAAATAAATAAAGCGCATGACCAACAGAT 587
QY 602 TCCAAAGGCGAGTCTCTTCGNGCCACCGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 661
Db 586 TCCAAATAAAGCGTAACTCTTCCCAACACCGCACTGCTGCTGCTGCTGCTGCTGCTGCT 527
QY 662 AGACTCTAGTGTGGCGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 721
Db 526 AAACCTCAACTATAACCTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACT 467
QY 722 CCCTGGCGCAGCTCTCTAGAGCGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 781
Db 466 CCCTAACGCACTCTCTTAACAGAAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 407
QY 782 AGCGCGGAGCGCGCGCGGAAAGCGCGCGGAGGGGTCTGTCACCGCGCGCGCGCGCGCG 841
Db 406 AACGGAACGACGCGGAAAGAAACGCGGAAAGAAATCTTACACCGCGCGCGCGCGCG 347
QY 842 CTGCTCTCGTCCCGCGCGCGCACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 901
Db 346 CTACCTCGTCCCGCGCGCGCACTTAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 287
QY 902 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 961
Db 286 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 227
QY 962 TGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1021
Db 226 TACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 167
QY 1022 ACCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1081
Db 166 ACCTAATAAATAATCAACATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 107
QY 1082 CCGGCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1141
Db 106 CCGACCTCTATCTGATTCGACATCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6
QY 1142 GCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1182
Db 46 ACCCTACGAAACGAATAACGAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 6
```

```
AX348476
AX348476.1 GI:18614512
.
synthetic construct.
SOURCE      synthetic construct
ORGANISM    artificial sequences.
REFERENCE   1
AUTHORS     Olek.A., Piepenbrock,C. and Berlin,K.
TITLE        Method and nucleic acids for pharmacogenomic methylation analysis
JOURNAL      Patent: WO 0202806-A 171 10-JAN-2002;
              Epigenomics AG (DE)
FEATURES    Location/Qualifiers
             source
               1..1186
               /organism="synthetic construct"
               /db_xref="taxon:32630"
               /note="chemically treated genomic DNA (Homo sapiens)"
BASE COUNT  224 a   98 c  347 g   517 t
ORIGIN
Query Match      59.8%; Score 710; DB 6; Length 1186;
Best Local Similarity 75.1%; Pred. No. 1.4e-118;
Matches 892; Conservative 0; Mismatches 294; Indels 1; Gaps 1;
QY 1 GATCGCGCACTGCACTCCAGCCTGGTGCAGAGCGAGACTCTGTCTCAAAAAAAA 60
Db 1 GATCGCGTATTGTTATTTTGTAGTTGGTGCAGAGCGAGATTTGTTTAAAAAAA 60
QY 61 AAAAAACCGCGAGGCTCAAAACAAAAACCTCGGAAAGCCCTGCGGCTCTTTTTTT 120
Db 61 AAAAAAGATCGTTAGGTTTAAATAAAAAATTCGGAAAAAGTTTGGCGGTTTTTTT 120
QY 121 TTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180
Db 121 TTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180
QY 181 CGSACTTGGCTCACTGCAACCTCTGCTCCAGGTTCAAGCAATCTTCTGCTCAGCC 240
Db 181 CGGATTTTGGTTTATGTTAATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 240
QY 241 TCCCAAGTAGCCACCAAGCGCCAGCTAATTTTGTANTTTTAGTAGAGACGGGGTTTCA 300
Db 241 TTTTAAAGTAGTTTATACGTTTGTAGTTAATTTTGTATTTTGTATTTAGTAGAGACGGGGTTTAT 300
QY 301 CATGTTGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
Db 301 TATGTTGTTAGTGTGGTTTGTGAATTTTGAATTTTGTAGTGTGATTTTATTCGTTTCGGTTTTT 360
QY 361 CAAAGTACTAGGATTAAGCGGTGAGCCACCGGTTCCAGCGCTGCGGCTGTTTAAATCA 420
Db 361 TAAAGTATTTAGGATTTATAGCGGTGAGTTATCGCGTTTACCGCTTTTGGCGGTTTTTAAATTA 420
QY 421 AGTAGAAAGCTGCAATATATACCACTTCTGCTTCGTTGCTTCAAGTGAAGCAAGAAATGG 480
Db 421 AGTAGAAAGCTGCAATATATATTTTGTTCGTTG-TTTTAGTAGAGAAAGAAATGG 479
QY 481 AAATGCAATCMTCTTATTTAGTTGTAGGAAACAGATCTCAAAACAGCAGTGTGTGTGNGACAAG 540
Db 480 AAATGTAATTTTATTTTATTTAGTTGTAGGAAATAGATTTTAAATAGTAGTTTGTGTGATAAG 539
QY 541 ACCGCAAGGAAACGTTGGGAACCTGTGCTGTGGCTTAGAGAGGCGCGGTGCAACAGACGG 600
Db 540 ATCGTAGGAAACGTTGGGAATTTGTTGTTGTTTGTAGAGAGGCGCGGTGCAATAGACGG 599
QY 601 TTCCCAAGGCGAGTCTCTCCCGCCACCGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Db 600 TTTTAAAGGCGTAGTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 659
QY 661 AAGACTCTCAGCTGTGGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCT 720
Db 660 AAGATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 719
QY 721 CCCCTGCGCAGCTCTCTTAGAGCGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 780
```


b	720	TTTTTGGCGTAGCTTTTATAGACGCGGGGTGGTCGCATTTCCGTCAGTAGAAGAGCGC	779
y	781	GAGCGCGGAGCGCGCGGGAATAAGCGCGCGGAGGGCTCTGCCACCGCGCCACTTTGG	840
b	780	GAGCGCGGAGCGGTGCGCGGAAAAGCGCGGAAAGGGGTTTTTGTTATCGCGTTATTGG	839
y	841	CCTGCTCCGTCGCGCGCGCACTTGGCCTCGCTCCGTCGCCCGCCGCCACTTCGCGCTG	900
b	840	TTTGTGTTTCGTTTCGTGCGGTATTTGGTTTGTTCGTTTCGCGTTATTTCGTTTG	899
y	901	CCTCGCTCCC CGCGCGCCCATGCTGTGGCCGCGCTCGGAGTTCGCGCGCGGCC	960
b	900	TTTTCGTTTTCGTTTCGTGCGGTATTTGGTTTGTTCGTTTCGCGTTGTGCGTCCGTTT	959
y	961	TTGCCCCCGCGCACAGGAGCGGAGCGCGGAGCGCGTCCGCGCACGCGGAGCTGCAG	1020
b	960	TTGTTTTTCGTGCTATAGGAGCGGAGTGCAGTTCGCGTTCGTCGAGTTGTGCGTCTGTA	1019
y	1021	TACCTGGGCGAGATCCAACAATCTCCGCTCGCGCGTCAAGAAGAACGACCGACGCGGC	1080
b	1020	TATTTGGGTAGATTTAATATATTTTTCGTTTCGCGGTAGGAAGACGATCGTACGGGT	1079
y	1081	ACGCGACCCCTGTCGTTATTTGGCATGACGCGCGCTACAGCTGAGAGGTGACCGCGCG	1140
b	1080	ATCGGTATTTTTCGCTATTCGCTATATAGGCGCGTTATATGTTTGAGAGGTACGTCGCG	1139
y	1141	GGCCCTCGGAGCGGGTGGCGGAGGAGGAGCGCGCGCTGGGGA	1187
b	1140	GGTTTTGCGGACGGGTGGCGGAGGAGGAGCGCGGTGGGGA	1186

PAT 14-FEB-2001

RESULT 14	
AR103889	Sequence 1 from patent US 6087489.
DEFINITION	Antisense oligonucleotide modulation of human thymidylate synthase expression
ACCESSION	AR103889
VERSION	AR103889.1 GI:12815477
KEYWORDS	Unknown.
SOURCE	Unknown.
ORGANISM	Unclassified.
REFERENCE	Dean,N.M.
AUTHORS	Antisense oligonucleotide modulation of human thymidylate synthase expression
TITLE	Patent: US 6087489-A 11-JUL-2000;
JOURNAL	Location/Qualifiers
FEATURES	1..1536 /source /organism="unknown"
BASE COUNT	390 a 369 c 399 g 378 t
ORIGIN	
Query Match	25.2%; Score 299.4; DB 6; Length 1536;
Best Local Similarity	99.7%; Pred. No. 1.8e-44;
Matches 300; Conservative	0; Mismatches 1; Indels 0; Gaps 0;

QY	833	CCACTTGGCTGCTCCGTCGCGCGCCACTTGGCTTCGCTCCGTCGCCCGCGCCAC	892
DB	14	CCACTTGGCTGCTCCGTCGCGCGCGCCACTTGGCTTCGCTCCGTCGCCCGCGCCAC	73
QY	893	TTGCGCTGCTCCTCGTCCCGCGCGCGCCATGCTGTGGCCGCTCGAGTTCGCC	952
DB	74	TTGCGCTGCTCCTCGTCCCGCGCGCGCCATGCTGTGGCCGCTCGAGTTCGCC	133
QY	953	GCGGCGCTTGGCCCCCGCGCACAGGAGCGGAGCGCGAGCGCGTCCGCGCACGGG	1012
DB	134	GCGGCGCTTGGCCCCCGCGCACAGGAGCGGAGCGCGAGCGCGTCCGCGCACGGG	193
QY	1013	AGCTGCACTACCTGGGCGAGATCCAACAATCTTCGCTCGCGCGTCAAGAGACACC	1072
DB	194	AGCTGCACTACCTGGGCGAGATCCAACAATCTTCGCTCGCGCGTCAAGAGACACC	253
QY	1073	GCACGGCGACCGCGACCTGTGCGTATTTCCGATGACGAGCGCGCTACACCTGAGGTG	1132

Db 254 GCACGGGACCGGACCCCTGTCGGTATTTCGGCATGCAGCGCGCTACAGCCTGAGAGATG 313

Qy 1133 A 1133

Db 314 A 314

RESULT 15
 AX397715 1536 bp DNA linear PAT 18-MAY-2002
 LOCUS Sequence 141 from Patent WO0210366.
 DEFINITION Sequence 141 from Patent WO0210366.
 ACCESSION AX397715
 VERSION AX397715.1 GI:21068437
 KEYWORDS human.
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
 Chen, H., Freimer, N. B. and Novak, T.
 Methods and compositions for diagnosing and treating chromosome-18P
 related disorders
 Patent: WO 0210366-A 141 07-FEB-2002;
 JOURNAL Millennium Pharmaceuticals, Inc. (US) ; The Regents of The
 University of California (US)
 FEATURES
 Location/Qualifiers
 1..1536
 /organism="Homo sapiens"
 /db_xref="taxon:9606" 378 t
 BASE COUNT 390 a 369 c 399 g 378 t
 ORIGIN

Query Match 25.2%; Score 299.4; DB 6; Length 1536;
 Best Local Similarity 99.7%; Pred. No. 1.8e-44;
 Matches 300; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 833 CCACTTGGCTGCTCCGTCGTCGCGCGCACCTTGGGCTGCCTTCCTCCGCGCGCCAC 892

Db 14 CCACTTGGCTGCTCCGTCGTCGCGCGCACCTTGGGCTGCCTTCCTCCGCGCGCCAC 73

Qy 893 TTGCGCTGCTCCGTCGTCGCGCGCGCGCGCATGCTGTGGCGGCTCGAGCTGCGC 952

Db 74 TTGCGCTGCTCCGTCGTCGCGCGCGCGCGCATGCTGTGGCGGCTCGAGCTGCGC 133

Qy 953 GCGGGCGCTTGGTCCCGCGCGCGCAGAGCGGAGCGCGCGCTCCGCGCGCACGGGG 1012

Db 134 GCGGGCGCTTGGTCCCGCGCGCGCAGAGCGGAGCGCGCGCTCCGCGCGCACGGGG 193

Qy 1013 AGCTGAGTACTTGGGGGAGATCCACACATCTCCGCTGGCGGCTCAGGAAGGACGACC 1072

Db 194 AGCTGAGTACTTGGGGGAGATCCACACATCTCCGCTGGCGGCTCAGGAAGGACGACC 253

Qy 1073 GCAGGGCACCGGACCTGTCGCGGTATTCGGCATCGAGCGCGCTACAGCCTGAGAGTG 1132

Db 254 GCACGGGACCGGACCTGTCGGTATTCGGCATCGAGCGCGCTACAGCCTGAGAGATG 313

Qy 1133 A 1133

Db 314 A 314

Search completed: November 28, 2002, 21:45:16
 Job time : 331.72 secs

Search completed: November 28, 2002, 21:45:16
Job time : 3331.72 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 29, 2002, 00:16:50 ; Search time 1562.32 Seconds

(without alignments)
12304.774 Million cell updates/sec

Title: US-09-963-333-7

Perfect score: 1187

Sequence: 1 gatcgccactgcactcca.....gagggagcgcgctggggga 1187

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues

Word size: 15

Total number of hits satisfying chosen parameters: 3111682

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database:

EST:

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vit:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Préd. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	335	28.2	1060	13	BM456638	AGENCOURT
2	329	27.7	741	9	AU118644	AU118644
3	318	26.8	841	14	BQ434233	BQ434233 AGENCOURT
4	309	26.0	1021	14	BQ685120	BQ685120 AGENCOURT
5	303	25.5	938	9	AL551990	AL551990
6	299	25.2	668	9	AL541351	AL541351

7	296	24.9	523	14	BM763658	BM763658 K-EST0045
8	295	24.9	738	9	AU130205	AU130205 AGENCOURT
9	280	23.6	948	14	BQ68205	BQ68205 AGENCOURT
10	278	23.4	739	9	AU143295	AU143295 AGENCOURT
11	278	23.4	819	9	AU143180	AU143180 AGENCOURT
12	278	23.4	1045	13	BM478045	BM478045 AGENCOURT
13	276	23.3	885	14	BQ424746	BQ424746 AGENCOURT
14	274	23.1	678	12	BG337663	BG337663 602435223
15	274	23.1	832	12	BG685317	BG685317 602141676
16	274	23.1	1055	14	BQ886916	BQ886916 AGENCOURT
17	272	22.9	970	14	BQ28426	BQ28426 AGENCOURT
18	266	22.4	994	14	BM904778	BM904778 AGENCOURT
19	265	22.3	994	9	AU117646	AU117646 AGENCOURT
20	265	22.3	1018	14	BM800322	BM800322 AGENCOURT
21	263	22.2	672	12	BG574267	BG574267 602596171
22	263	22.2	963	12	BG337540	BG337540 602435078
23	263	22.2	985	13	BM466216	BM466216 AGENCOURT
24	262	22.1	406	14	BM834317	BM834317 K-EST0109
25	262	22.1	688	9	AL543217	AL543217 AL543217
26	262	22.1	711	9	AU142980	AU142980 AGENCOURT
27	262	22.1	888	13	BM413631	BM413631 AGENCOURT
28	262	22.1	889	14	BQ420981	BQ420981 AGENCOURT
29	262	22.1	905	13	B1597722	B1597722 603248376
30	262	22.1	1020	9	AL547612	AL547612 AL547612
31	260	21.9	529	12	BE794349	BE794349 601589526
32	258	21.7	1533	11	BC018858	BC018858 Homo sapi
33	254	21.4	442	10	BE280876	BE280876 601155452
34	254	21.4	465	12	BG749145	BG749145 602706761
35	251	21.1	904	14	BQ642214	BQ642214 AGENCOURT
36	251	21.1	972	14	BQ642223	BQ642223 AGENCOURT
37	251	21.1	984	14	BQ061952	BQ061952 AGENCOURT
38	251	21.1	1027	14	BQ065204	BQ065204 AGENCOURT
39	251	21.1	1048	14	BQ065413	BQ065413 AGENCOURT
40	251	21.1	1071	14	BQ057866	BQ057866 AGENCOURT
41	251	21.1	1072	14	BQ066271	BQ066271 AGENCOURT
42	251	21.1	1084	14	BQ063450	BQ063450 AGENCOURT
43	251	21.1	1135	14	BQ062012	BQ062012 AGENCOURT
44	251	21.1	1549	14	BQ073452	BQ073452 AGENCOURT
45	251	21.1	1614	14	BQ057801	BQ057801 AGENCOURT

ALIGNMENTS

RESULT 1
BM456638
LOCUS
DEFINITION
AGENCOURT_6408748 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:5496315
5', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
human.
EST.
BM456638.1 GI:18505678

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

NIH-MGC <http://mgs.nci.nih.gov/>
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov

Tissue Procurement: Lou Staudt
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.G.E. Consortium/LLNL at:
<http://image.llnl.gov>

Plate: LLAM12124 row: k column: 04
High quality sequence stop: 530.
Location/Qualifiers
1..1060

FEATURES
source

```
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5496315"
/clone_lib="NIH_MGC_85"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: lymph; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.867 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC library."
BASE COUNT      217 a      396 c      259 g      188 t
ORIGIN

Query Match      28.2%; Score 335; DB 13; Length 1060;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 795 CCGGGAAGAAGCGCGGAGGGTCTGCTCCACCGCGCCACTTGGCTGCTCCGTCCTCC 854
Db 1 CCGGGAAGAAGCGCGGAGGGTCTGCTCCACCGCGCCACTTGGCTGCTCCGTCCTCC 60

QY 855 GCGCGCCACTTGGCTGCTCCGTCCTCCGCGCGCGCCACTTGGCTGCTCCGTCCTCC 914
Db 61 GCGCGCCACTTGGCTGCTCCGTCCTCCGCGCGCGCCACTTGGCTGCTCCGTCCTCC 120

QY 915 CCGCGCGCCACTTGGCTGCTCCGTCCTCCGCGCGCGCCACTTGGCTGCTCCGTCCTCC 974
Db 121 CCGCGCGCCACTTGGCTGCTCCGTCCTCCGCGCGCGCCACTTGGCTGCTCCGTCCTCC 180

QY 975 ACAGAGCGGAGCGCGGAGCGCTGCGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCG 1034
Db 181 ACAGAGCGGAGCGCGGAGCGCTGCGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCG 240

QY 1035 CCAACACATCTCCGTCGCGGCTCAGAGAGGACGCGCGCGCGCGCGCGCGCGCGCGCG 1094
Db 241 CCAACACATCTCCGTCGCGGCTCAGAGAGGACGCGCGCGCGCGCGCGCGCGCGCGCG 300

QY 1095 GGTATTGGGATCGAGGCGGCTACAGCTGAGAG 1129
Db 301 GGTATTGGGATCGAGGCGGCTACAGCTGAGAG 335

RESULT 2
AUI18644
LOCUS      AUI18644 HEMBA1 Homo sapiens cDNA clone HEMBA1004057 5', mRNA
DEFINITION      sequence.
ACCESSION      AUI18644.1 GI:109333771
VERSION        AUI18644.1
KEYWORDS       EST.
SOURCE         human.
ORGANISM       Homo sapiens
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS       Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE         1 (bases 1 to 741)
JOURNAL       Ota.T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
COMMENT       Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
              Isogai,T.
              HRI human cDNA project
              Unpublished (2000)
              Contact: Takao Isogai
              Genomics Laboratory
              Helix Research Institute
              1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
              Tel: 81-438-52-3975
              Fax: 81-438-52-3986
              Email: genomics@hri.co.jp
              HRI human cDNA project, 5'- & 3'-end one pass sequencing: Helix
              Research Institute; cDNA library construction: Department of
              Virology, Institute of Medical Science, University of Tokyo, and
              Helix Research Institute.
              Location/Qualifiers
                1..741
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:5496315"
                /clone_lib="NIH_MGC_85"
                /tissue_type="lymphoma, cell line"
                /lab_host="DH10B (phage-resistant)"
                /notes="Organ: lymph; Vector: pCMV-SPORT6; Site 1: NotI;
                Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
                Average insert size 1.867 kb. Library enriched for
                full-length clones and constructed by Life Technologies.
                Note: this is a NIH_MGC library."
                BASE COUNT      217 a      396 c      259 g      188 t
                ORIGIN

Source Match      27.7%; Score 329; DB 9; Length 741;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 329; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 801 AAAAGGCGCGGAAAGGGTCTGCTCCACCGCGCCACTTGGCTGCTCCGTCCTCCGTCCT 860
Db 2 AAAAGGCGCGGAAAGGGTCTGCTCCACCGCGCCACTTGGCTGCTCCGTCCTCCGTCCT 61

QY 861 CCATTGGCTGCTGCTCCGTCCTCCGCGCGCGCCACTTGGCTGCTCCGTCCTCCGTCCT 920
Db 62 CCATTGGCTGCTGCTCCGTCCTCCGCGCGCGCCACTTGGCTGCTCCGTCCTCCGTCCT 121

QY 921 CGCCATGCTGTGGCGGCTCGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 980
Db 122 CGCCATGCTGTGGCGGCTCGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 181

QY 981 GCGGAGCGCGGAGCGCGCTGCGAGAGGACGCGCGCGCGCGCGCGCGCGCGCGCGCG 1040
Db 182 GCGGAGCGCGGAGCGCGCTGCGAGAGGACGCGCGCGCGCGCGCGCGCGCGCGCGCG 241

QY 1041 CATCTTCGCTGCGGCTCAGAGAGGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1100
Db 242 CATCTTCGCTGCGGCTCAGAGAGGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 301

QY 1101 CGGCATGCGAGGCGGCTACAGCTGAGAG 1129
Db 302 CGGCATGCGAGGCGGCTACAGCTGAGAG 330

RESULT 3
BQ434233
LOCUS      BQ434233 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6159722
DEFINITION      5', mRNA sequence.
ACCESSION      BQ434233
VERSION        BQ434233.1 GI:21173309
KEYWORDS       EST.
SOURCE         human.
ORGANISM       Homo sapiens
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS       Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE         1 (bases 1 to 841)
JOURNAL       NIH-MGC http://mgc.nci.nih.gov/.
COMMENT       National Institutes of Health, Mammalian Gene Collection (MGC)
              Unpublished (1999)
              Contact: Robert Strausberg, Ph.D.
              Email: cgabbs@mail.nih.gov
              Tissue Procurement: ATCC/DCTP/DTF
              cDNA Library Preparation: Life Technologies, Inc.
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
              DNA Sequencing by: Agencourt Bioscience Corporation
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LIML at:
              http://image.llnl.gov
              Plate: LIML13509 row: e column: 03
              High quality sequence stop: 639.
              Location/Qualifiers
                1..841
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:6159722"
                /clone_lib="NIH_MGC_72"
                /tissue_type="whole embryo, mainly head"
                /dev_stage="embryo, 10 weeks"
                /note="vector: pME18SFL3"
                BASE COUNT      153 a      227 c      218 g      140 t
                3 others
                ORIGIN

Query Match      27.7%; Score 329; DB 9; Length 741;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 329; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 801 AAAAGGCGCGGAAAGGGTCTGCTCCACCGCGCCACTTGGCTGCTCCGTCCTCCGTCCT 860
Db 2 AAAAGGCGCGGAAAGGGTCTGCTCCACCGCGCCACTTGGCTGCTCCGTCCTCCGTCCT 61

QY 861 CCATTGGCTGCTGCTCCGTCCTCCGCGCGCGCCACTTGGCTGCTCCGTCCTCCGTCCT 920
Db 62 CCATTGGCTGCTGCTCCGTCCTCCGCGCGCGCCACTTGGCTGCTCCGTCCTCCGTCCT 121

QY 921 CGCCATGCTGTGGCGGCTCGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 980
Db 122 CGCCATGCTGTGGCGGCTCGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 181

QY 981 GCGGAGCGCGGAGCGCGCTGCGAGAGGACGCGCGCGCGCGCGCGCGCGCGCGCGCG 1040
Db 182 GCGGAGCGCGGAGCGCGCTGCGAGAGGACGCGCGCGCGCGCGCGCGCGCGCGCGCG 241

QY 1041 CATCTTCGCTGCGGCTCAGAGAGGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1100
Db 242 CATCTTCGCTGCGGCTCAGAGAGGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 301

QY 1101 CGGCATGCGAGGCGGCTACAGCTGAGAG 1129
Db 302 CGGCATGCGAGGCGGCTACAGCTGAGAG 330

RESULT 3
BQ434233
LOCUS      BQ434233 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6159722
DEFINITION      5', mRNA sequence.
ACCESSION      BQ434233
VERSION        BQ434233.1 GI:21173309
KEYWORDS       EST.
SOURCE         human.
ORGANISM       Homo sapiens
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS       Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE         1 (bases 1 to 841)
JOURNAL       NIH-MGC http://mgc.nci.nih.gov/.
COMMENT       National Institutes of Health, Mammalian Gene Collection (MGC)
              Unpublished (1999)
              Contact: Robert Strausberg, Ph.D.
              Email: cgabbs@mail.nih.gov
              Tissue Procurement: ATCC/DCTP/DTF
              cDNA Library Preparation: Life Technologies, Inc.
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
              DNA Sequencing by: Agencourt Bioscience Corporation
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LIML at:
              http://image.llnl.gov
              Plate: LIML13509 row: e column: 03
              High quality sequence stop: 639.
              Location/Qualifiers
                1..841
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:6159722"
                /clone_lib="NIH_MGC_72"
                /tissue_type="whole embryo, mainly head"
                /dev_stage="embryo, 10 weeks"
                /note="vector: pME18SFL3"
                BASE COUNT      153 a      227 c      218 g      140 t
                3 others
                ORIGIN
```

/lab_host="DH10B (phage-resistant)"
 /note="Organ: pancreas; Vector: pOTB7; Site 1: XhoI;
 Site 2: EcoRI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCACGAG(G). Library constructed by
 Ling Hong in the laboratory of Gerald M. Rubin (University
 of California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library."
 202 a 335 c 289 g 194 t 1 others

BASE COUNT 202 a 335 c 289 g 194 t 1 others
 ORIGIN
 Query Match 26.0%; Score 309; DB 14; Length 1021;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 309; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 821 CTGTCACGCGCCACTTGGCTCGCTCCGCGCGCGCCACTTGGCTCGCTCGCTCGCT 880
 Db 11 CTGTCACGCGCCACTTGGCTCGCTCCGCGCGCGCCACTTGGCTCGCTCGCTCGCT 70
 QY 881 CGCGCGCGCCACTTGGCTCGCTCCGCGCGCGCCACTTGGCTCGCTCGCTCGCTCGCT 940
 Db 71 CGCGCGCGCCACTTGGCTCGCTCCGCGCGCGCCACTTGGCTCGCTCGCTCGCTCGCT 130
 QY 941 CGGAGCTGCGCGCGCGCTTGGCTCGCTCCGCGCGCGCCACTTGGCTCGCTCGCTCGCT 1000
 Db 131 CGGAGCTGCGCGCGCGCTTGGCTCGCTCCGCGCGCGCCACTTGGCTCGCTCGCTCGCT 190
 QY 1001 CGCGCGCGCCACTTGGCTCGCTCCGCGCGCGCCACTTGGCTCGCTCGCTCGCTCGCT 1060
 Db 191 CGCGCGCGCCACTTGGCTCGCTCCGCGCGCGCCACTTGGCTCGCTCGCTCGCTCGCT 250
 QY 1061 GGAAGGACGACGCGCGCGCTTGGCTCGCTCCGCGCGCGCCACTTGGCTCGCTCGCTCGCT 1120
 Db 251 GGAAGGACGACGCGCGCGCTTGGCTCGCTCCGCGCGCGCCACTTGGCTCGCTCGCTCGCT 310
 QY 1121 GCTGTGAGAG 1129
 Db 311 GCTGTGAGAG 319

RESULT 5
 AL551990 938 bp mRNA linear EST 16-FEB-2001
 LOCUS AL551990 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0D1060YE04 5
 DEFINITION prime, mRNA sequence.
 ACCESSION AL551990.1 GI:12890459
 VERSION AL551990.1
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Li W.B., Gruber, C., Jessee, J., and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
 Location/Qualifiers

FEATURES
 source
 1..938
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CS0D1060YE04"
 /clone_lib="LTI_NFL006_PL2"
 /tissue_type="placenta"
 /note="Vector: pCWSport 6; Site 1: NotI; 1st strand cDNA
 was primed with a NotI-oligo(dT) primer. Five prime end
 enriched, double-stranded cDNA was digested with Not I and
 cloned into the Not I and Eco RV sites of the pCWSport 6
 vector. Library was normalized. Library was constructed by

/tissue_type="melanotic melanoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: skin; Vector: pCW-Sport6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 2 kb. Library constructed by Life
 Technologies."
 176 a 263 c 241 g 161 t

BASE COUNT 176 a 263 c 241 g 161 t
 ORIGIN
 Query Match 26.8%; Score 318; DB 14; Length 841;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 318; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 812 GGAAGGAGTCTGTCACGCGCCACTTGGCTCGCTCCGCGCGCGCCACTTGGCTCGCT 871
 Db 1 GGAAGGAGTCTGTCACGCGCCACTTGGCTCGCTCCGCGCGCGCCACTTGGCTCGCT 60
 QY 872 GCTTCGCTCGCGCGCGCCACTTGGCTCGCTCCGCGCGCGCCACTTGGCTCGCTCGCT 931
 Db 61 GCTTCGCTCGCGCGCGCCACTTGGCTCGCTCCGCGCGCGCCACTTGGCTCGCTCGCT 120
 QY 932 TGGCGCGCTGCGAGCTGCGCGCGCGCTTGGCTCGCTCCGCGCGCGCCACTTGGCTCGCT 991
 Db 121 TGGCGCGCTGCGAGCTGCGCGCGCGCTTGGCTCGCTCCGCGCGCGCCACTTGGCTCGCT 180
 QY 992 AGCGCGCTGCGCGCGCGAGCTGCGAGTACCTGGGCGAGATCCACACATCTCCGCT 1051
 Db 181 AGCGCGCTGCGCGCGCGAGCTGCGAGTACCTGGGCGAGATCCACACATCTCCGCT 240
 QY 1052 GCGCGCTCAGGAAGGACGCGCGCGCGCTTGGCTCGCTCCGCGCGCGCCACTTGGCTCGCT 1111
 Db 241 GCGCGCTCAGGAAGGACGCGCGCGCGCTTGGCTCGCTCCGCGCGCGCCACTTGGCTCGCT 300
 QY 1112 CGCGCTACAGCTCGAGAG 1129
 Db 301 CGCGCTACAGCTCGAGAG 318

RESULT 4
 BQ685120 1021 bp mRNA linear EST 15-JUL-2002
 LOCUS AGENCOURT 8344472 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6251301
 DEFINITION 5', mRNA sequence.

ACCESSION BQ685120
 VERSION BQ685120.1 GI:21810436
 KEYWORDS EST.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: c9apb8-remail.nih.gov
 Tissue Procurement: ATCC

CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LLCW2395 row: 1 column: 22
 High quality sequence start: 12
 High quality sequence stop: 409.
 Location/Qualifiers

FEATURES
 source
 1..1021
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:6251301"
 /clone_lib="NIH_MGC_110"
 /tissue_type="ductal carcinoma, cell line"

Life Technologies. Contact : Peng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 202 a 288 c 259 g 188 t 1 others
ORIGIN

Query Match 25.5%; Score 303; DB 9; Length 938;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 303; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 827 ACCGCGCCACTTGGCTCCCTCCGCGCGCCACTTGGCTCCGCTCCGTCGCCGCG 886

Db 1 ACCGCGCCACTTGGCTCCCTCCGCGCGCCACTTGGCTCCGCTCCGTCGCCGCG 60

QY 887 CGCCACTTGGCTCCCTCCGCGCGCCACTTGGCTCCGCTCCGTCGCCGCG 946

Db 61 CGCCACTTGGCTCCCTCCGCGCGCCACTTGGCTCCGCTCCGTCGCCGCG 120

QY 947 TGCCTGCGCGCCCTTGCCTCCGCGCGCCACTTGGCTCCGCTCCGTCGCCGCG 1006

Db 121 TGCCTGCGCGCCCTTGCCTCCGCGCGCCACTTGGCTCCGCTCCGTCGCCGCG 180

QY 1007 ACGGAGCTGAGTACCTGGGGGAGATCCACACATCTCCGTCGGGCGTCAGGAAG 1066

Db 181 ACGGAGCTGAGTACCTGGGGGAGATCCACACATCTCCGTCGGGCGTCAGGAAG 240

QY 1067 ACGACGCGACGGGCGCGCCCTGCGGTATTCGGCATGCGGCGGCTACAGCTGA 1126

Db 241 ACGACGCGACGGGCGCGCCCTGCGGTATTCGGCATGCGGCGGCTACAGCTGA 300

QY 1127 GAG 1129

Db 301 GAG 303

RESULT 6

AL541351 668 bp mRNA linear EST 16-FEB-2001
LOCUS
DEFINITION AL541351 LTI FL002_P1 Homo sapiens cDNA clone CS0D006YA08 5 prime
mRNA sequence.

ACCESSION AL541351

VERSION AL541351.1 GI:12872339

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 668)

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished (2001)

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

Location/Qualifiers

1. 668

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="CS0D006YA08"

/lab_host="VDH10B"

/note="Organ: placenta; Vector: pCMVSPORT 6; 1st strand

cDNA was primed with a NotI-oligo(dT) primer. Five prime

end enriched, double-stranded cDNA was digested with Not I

and cloned into the Not I and Eco RV sites of the

pCMVSPORT 6 vector. Library was constructed by Life

Technologies. Contact : Peng Liang Life Technologies, a

division of Invitrogen 9800 Medical Center Drive Rockville

, Maryland 20850, USA Fax : (1) 301 610 8371 Email :

fliang@lifetech.com URL :

BASE COUNT 139 a 208 c 192 g 126 t 3 others

ORIGIN

Query Match 25.2%; Score 299; DB 9; Length 668;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 299; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 820 TCCTGCCACGCGCCACTTGGCTCCCTCCGCGCGCCACTTGGCTCCGCTCCGTCGCCGCG 879

Db 1 TCCTGCCACGCGCCACTTGGCTCCCTCCGCGCGCCACTTGGCTCCGCTCCGTCGCCGCG 60

QY 880 CCCGCGCGCGCTTGCCTCCGCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 939

Db 61 CCCGCGCGCGCTTGCCTCCGCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120

QY 940 TCGAGAGTCCG 999

Db 121 TCGAGAGTCCG 180

QY 1000 CCGCGCGACGCGGAGCTGAGTACCTGGGCGAGATCCAAACATCTCCGTCGGGCGGTC 1059

Db 181 CCGCGCGACGCGGAGCTGAGTACCTGGGCGAGATCCAAACATCTCCGTCGGGCGGTC 240

QY 1060 AGGAAGGACGACGCGACCGGCGACCGGCGACCGGCGACCGGCGACCGGCGCGCTA 1118

Db 241 AGGAAGGACGACGCGACCGGCGACCGGCGACCGGCGACCGGCGACCGGCGCGCTA 299

RESULT 7

BM763658

LOCUS

DEFINITION K-EST0045041 S13KMS5 Homo sapiens cDNA clone S13KMS5-14-F09 5',

mRNA sequence.

ACCESSION BM763658

VERSION BM763658.1 GI:19093273

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 523)

Kim, N.S., Hahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R.,

Oh, K.J., Cheong, J.E., Sohn, H.Y., Kim, J.M., Park, H.S., Kim, S. and

Kim, Y.S.

21C Frontier Korean EST Project 2001

Unpublished (2002)

Contact: Kim YS

Genome Research Center

Korea Research Institute of Bioscience & Biotechnology

52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea

Tel: +82-42-860-4470

Fax: +82-42-860-4409

Email: yongsung@mail.kribb.re.kr

Plate: 14 row: F column: 09

High quality sequence stop: 523.

Location/Qualifiers

1. 523

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="S13KMS5-14-F09"

/clone_lib="S13KMS5"

/tissue_type="myeloma"

/lab_host="Top10P"

/note="Vector: pcNS; Site 1: EcoRI; Site 2: NotI; The poly

(A) + RNA was dephosphorylated with bacterial alkaline

phosphatase (BAP) and then decapped with tobacco acid

pyrophosphatase (TAP). The decapped intact mRNA was

ligated with DNA-RNA linker including EcoR I site by

treatment of T4 RNA ligase and the first strand cDNA was

synthesized from oligo dT-selected mRNA by priming with

dT-tailed vector. The dT-tailed vector was adjusted to

have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10⁺ by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

```
BASE COUNT      89 a   173 c   163 g   98 t
ORIGIN
Query Match      24.9%; Score 296; DB 14; Length 523;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 834 CACTTGGCTGCTCCGCTCCGCGCGGCCACTTGGCTTGCCTCCGTCGCCCGCGCCACT 893
Db 2 CACTTGGCTGCTCCGCTCCGCGCGGCCACTTGGCTTGCCTCCGTCGCCCGCGCCACT 61

QY 894 TCGCTTGCCTCCGTCGCCCGCGCGCCACTTGGCTTGCCTCCGTCGCCCGCGCGCG 953
Db 62 TCGCTTGCCTCCGTCGCCCGCGCGCCACTTGGCTTGCCTCCGTCGCCCGCGCGCG 121

QY 954 CCGGCTTGGCTCCGTCGCCCGCGCGCCACTTGGCTTGCCTCCGTCGCCCGCGCGCG 1013
Db 122 CCGGCTTGGCTCCGTCGCCCGCGCGCCACTTGGCTTGCCTCCGTCGCCCGCGCGCG 181

QY 1014 GTCGAGTACCTGGGCGGAGTCCAAACATCTCCGTCGCCCGCGCGCGCGCGCGCG 1073
Db 182 GTCGAGTACCTGGGCGGAGTCCAAACATCTCCGTCGCCCGCGCGCGCGCGCGCG 241

QY 1074 CACGGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1129
Db 242 CACGGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 297
```

```
RESULT 8
LOCUS      AU130205      738 bp      mRNA      linear      EST 01-AUG-2002
DEFINITION AU130205 NT2RP3 Homo sapiens cDNA clone NT2RP3000412 5', mRNA
sequence.
ACCESSION  AU130205
VERSION     AU130205.1 GI:10990559
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 738)
AUTHORS     Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
Isogai,T.
HRI human cDNA project
Unpublished (2000)
Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
Research Institute; cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.
FEATURES
source
1. .738
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="NT2RP3000412"
/clone_lib="NT2RP3"
/cell_type="teratocarcinoma"
/cell_line="NT2"
```

/note="Vector: pME189FL3; mRNA from NT2 neuronal precursor cells after 2-weeks retinoic acid (RA) induction"

```
BASE COUNT      164 a   216 c   212 g   143 t   3 others
ORIGIN
```

Query Match 24.9%; Score 295; DB 9; Length 738;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 295; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 835 ACTTGGCTTGCCTCCGCTCCGCGCGGCCACTTGGCTTGCCTCCGTCGCCCGCGCCACTT 894
Db 1 ACTTGGCTTGCCTCCGCTCCGCGCGGCCACTTGGCTTGCCTCCGTCGCCCGCGCCACTT 60

QY 895 CGCTTGCCTCCGTCGCCCGCGCGGCCACTTGGCTTGCCTCCGTCGCCCGCGCGCG 954
Db 61 CGCTTGCCTCCGTCGCCCGCGCGGCCACTTGGCTTGCCTCCGTCGCCCGCGCGCG 120

QY 955 CGGCTTGCCTCCGTCGCCCGCGCGGCCACTTGGCTTGCCTCCGTCGCCCGCGCGCG 1014
Db 121 CGGCTTGCCTCCGTCGCCCGCGCGGCCACTTGGCTTGCCTCCGTCGCCCGCGCGCG 180

QY 1015 CTGCACTTACTGGGCGGAGTCCAAACATCTCCGTCGCCCGCGCGCGCGCGCGCGCG 1074
Db 181 CTGCACTTACTGGGCGGAGTCCAAACATCTCCGTCGCCCGCGCGCGCGCGCGCGCG 240

QY 1075 ACGGCGCACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1129
Db 241 ACGGCGCACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 295
```

```
RESULT 9
LOCUS      BQ068205      948 bp      mRNA      linear      EST 02-APR-2002
DEFINITION AGENCOURT 6794358 NIH_MGC_121 Homo sapiens cDNA clone IMAGE:5770435
5', mRNA sequence.
ACCESSION  BQ068205
VERSION     BQ068205.1 GI:19897251
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

```
REFERENCE   1 (bases 1 to 948)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cygaps-remail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
```

```
Plate: L1AM12833 row: h column: 20
High quality sequence stop: 717.
Location/Qualifiers
1. .948
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5770435"
/clone_lib="NIH_MGC_121"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMV-SPORT6; Site: 1: NotI;
Site 2: EcoRV (destroyed); RNA source anonymous pool of 3
fetal brains, female age 20 weeks, female age 24 weeks,
and male age 26 weeks. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.7 kb, insert size range
0.7-3.5 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 017. Note:
```

```
FEATURES
source
```



```

Db 2 AAAAGCGCGGAGAGGGTCTTGCCACCGCGCCACTTGGCTCGCTCCGTCCCGCCGCG 61
Qy 861 CCACTTGGCTGCTCCCTCCGCGCGCGCCACTTGGCTCGCTCCGTCCCGCCGCGCG 920
Db 62 CCACTTGGCTGCTCCGTCCGCGCGCGCCACTTGGCTCGCTCCGTCCCGCCGCGCG 121
Qy 921 CGCATGCTGCTGGCGGCTCGGAGTGGCGCGCGCGCTTGGCGCGCGCGCGCGCA 980
Db 122 CGCATGCTGCTGGCGGCTCGGAGTGGCGCGCGCGCTTGGCGCGCGCGCGCGCA 181
Qy 981 GCGGAGCGCGAGCGCGGTCCGCGCGCGCGGAGCTGAGTACTCGGGCGCATCAACA 1040
Db 182 GCGGAGCGCGAGCGCGGTCCGCGCGCGCGGAGCTGAGTACTCGGGCGCATCAACA 241
Qy 1041 CATCTCTCGCTGCGCGCTCAGGAAGACGACCGCGCGCGCGCGCGCGCGCGTATT 1100
Db 242 CATCTCTCGCTGCGCGCTCAGGAAGACGACCGCGCGCGCGCGCGCGCGCGTATT 301
Qy 1101 CGGCATGCGAGCGCGCTTACAGCTGAGAG 1129
Db 302 CGGCATGCGAGCGCGCTTACAGCTGAGAG 330

RESULT 12
BM478045 1045 bp mRNA linear EST 05-FEB-2002
LOCUS BM478045
DEFINITION AGENCOURT 6482920 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:5552680
5', mRNA sequence.
ACCESSION BM478045
VERSION BM478045.1 GI:18527087
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1045)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue procurement: Lou Staudt
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12268 row: 9 column: 17
High quality sequence stop: 687.
Location/Qualifiers
1..1045
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5552680"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lymph; Vector: PCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dt primed.
Average insert size 1.867 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC library."
BASE COUNT 242 a 297 c 286 g 220 t
ORIGIN
Query Match 23.4%; Score 278; DB 13; Length 1045;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 278; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 852 CCGCGCGCGCACTTGGCTGCTCGGTCGCGCGCGCGCACTTGGCTGCTCGCTCCCG 911
Db 1 CCGCGCGCGCACTTGGCTGCTCGGTCGCGCGCGCGCACTTGGCTGCTCGCTCCCG 60

```

```

Qy 912 CGCCGCGCGCATGCTTGTGGCGCGGTTCGGAGCTGCGCGCGCGCCCTTGTCCCGCCGCG 971
Db 61 CGCCGCGCGCATGCTTGTGGCGCGGTTCGGAGCTGCGCGCGCGCCCTTGTCCCGCCGCG 120
Qy 972 CGCACAGGAGCGGAGCGCGAGCGCGTTCGGCGCGCGAGCGGGAGCTGAGTACTTGGGGCA 1031
Db 121 CGCACAGGAGCGGAGCGCGAGCGCGTTCGGCGCGCGAGCGGGAGCTGAGTACTTGGGGCA 180
Qy 1032 GATCCAAACACATCTCTCCGCTGCGCGCTCAGGAAGACGACCGCGCGCGCGCACCCCT 1091
Db 181 GATCCAAACACATCTCTCCGCTGCGCGCTCAGGAAGACGACCGCGCGCGCGCACCCCT 240
Qy 1092 GTCGATTCGGCATGACAGCGCGCTTACAGCTGAGAG 1129
Db 241 GTCGATTCGGCATGACAGCGCGCTTACAGCTGAGAG 278

RESULT 13
BM424746 885 bp mRNA linear EST 23-MAY-2002
LOCUS BM424746
DEFINITION AGENCOURT 7912639 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6150060
5', mRNA sequence.
ACCESSION BM424746
VERSION BM424746.1 GI:21120061
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 885)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue procurement: ATCC/DCTP/DTP
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13484 row: b column: 13
High quality sequence stop: 528.
Location/Qualifiers
1..885
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6150060"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: PCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: oligo dt.
Average insert size 2 kb. Library constructed by Life
Technologies."
BASE COUNT 177 a 287 c 257 g 164 t
ORIGIN
Query Match 23.3%; Score 276; DB 14; Length 885;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 854 CGCGCGCGCACTTGGCTGCTCGGTCGCGCGCGCGCACTTGGCTGCTCGCTCCCGCG 913
Db 48 CGCGCGCGCACTTGGCTGCTCGGTCGCGCGCGCGCACTTGGCTGCTCGCTCCCGCG 107
Qy 914 CGCGCGCGCGCACTTGGCTGCTCGGTCGCGCGCGCGCACTTGGCTGCTCGCTCCCGCG 973
Db 108 CGCGCGCGCGCACTTGGCTGCTCGGTCGCGCGCGCGCACTTGGCTGCTCGCTCCCGCG 167
Qy 974 CACAGGAGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 1033

```


QY 1036 CAACACATCTCGCTGGCGCTCAGGAAGGACGACCGCACGGGACACGGGACCCCTGTTCG 1095
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 199 CAACACATCTCTCCCTCGCGCTCAGGAAGGACGACCGCACGGGACCCCTGTTCG 258
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 1096 GTATTGGGATGCAAGCGCGCTACAGCCTTGAGAG 1129
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 259 GTATTGGGATGCAAGCGCGCTACAGCCTTGAGAG 292
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Search completed: November 29, 2002, 03:28:11
Job time : 1576.32 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

M nucleic - nucleic search, using sw model
Run on: November 29, 2002, 00:40:45 ; Search time 676.542 Seconds
(without alignments)
1187.536 Million cell updates/sec

Title: US-09-963-333-7
Perfect score: 1187
Sequence: 1 gatcgccactgcactcca.....gagggggcgcgcgtg999ga 1187.

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

searched: 344316 seqs, 338423730 residues

ord size : 15

total number of hits satisfying chosen parameters: 60992

inimum DB seq length: 0
aximum DB seq length: 2000000000

ost-processing: Listing first 45 summaries

- Database :
- 1: Pending Patents NA New.*
 - 2: /cgn2_6/ptodata/1/pna/PCT NEW COMB.seq.*
 - 3: /cgn2_6/ptodata/1/pna/US05 NEW COMB.seq.*
 - 4: /cgn2_6/ptodata/1/pna/US07 NEW COMB.seq.*
 - 5: /cgn2_6/ptodata/1/pna/US08 NEW COMB.seq.*
 - 6: /cgn2_6/ptodata/1/pna/US09 NEW COMB.seq.*
 - 7: /cgn2_6/ptodata/1/pna/US10 NEW COMB.seq.*
 - 8: /cgn2_6/ptodata/1/pna/US60 NEW_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1187	100.0	1187	5	US-09-658-659B-7
2	811	68.3	18597	5	US-09-658-659B-8
3	301	25.4	1942	5	US-09-724-676-34523
4	301	25.4	1942	5	US-09-724-676A-34523
5	301	25.4	2042	5	US-09-724-676-34515
6	301	25.4	2042	5	US-09-724-676A-34515
7	301	25.4	2391	5	US-09-724-676-34597
8	301	25.4	2391	5	US-09-724-676A-34597
9	301	25.4	2411	5	US-09-724-676-34535
10	301	25.4	2411	5	US-09-724-676A-34535
11	301	25.4	2491	5	US-09-724-676-34589
12	301	25.4	2491	5	US-09-724-676A-34589
13	301	25.4	2527	5	US-09-724-676-34548
14	301	25.4	2527	5	US-09-724-676A-34548
15	301	25.4	3023	5	US-09-724-676-34565
16	301	25.4	3023	5	US-09-724-676A-34565
17	301	25.4	3123	5	US-09-724-676-34557
18	301	25.4	3123	5	US-09-724-676A-34557
19	301	25.4	3139	5	US-09-724-676-34581
20	301	25.4	3139	5	US-09-724-676A-34581
21	301	25.4	3239	5	US-09-724-676-34573
22	301	25.4	3239	5	US-09-724-676A-34573
23	297	25.0	1336	5	US-09-658-659B-6
24	262	22.1	3298	6	US-10-240-965-200
25	111	9.4	2083	5	US-09-724-676-34524
26	111	9.4	2083	5	US-09-724-676A-34524

c 27	111	9.4	2183	5	US-09-724-676-34516	Sequence 34516, A
c 28	111	9.4	2183	5	US-09-724-676A-34516	Sequence 34516, A
c 29	111	9.4	2552	5	US-09-724-676-34536	Sequence 34536, A
c 30	111	9.4	2552	5	US-09-724-676A-34536	Sequence 34536, A
c 31	111	9.4	2632	5	US-09-724-676-34590	Sequence 34590, A
c 32	111	9.4	2632	5	US-09-724-676A-34590	Sequence 34590, A
c 33	111	9.4	2652	5	US-09-724-676-34529	Sequence 34529, A
c 34	111	9.4	2652	5	US-09-724-676A-34529	Sequence 34529, A
c 35	111	9.4	2668	5	US-09-724-676-34549	Sequence 34549, A
c 36	111	9.4	2668	5	US-09-724-676A-34549	Sequence 34549, A
c 37	111	9.4	2768	5	US-09-724-676-34542	Sequence 34542, A
c 38	111	9.4	2768	5	US-09-724-676A-34542	Sequence 34542, A
c 39	111	9.4	3164	5	US-09-724-676-34566	Sequence 34566, A
c 40	111	9.4	3164	5	US-09-724-676A-34566	Sequence 34566, A
c 41	111	9.4	3253	5	US-09-724-676-34525	Sequence 34525, A
c 42	111	9.4	3253	5	US-09-724-676A-34525	Sequence 34525, A
c 43	111	9.4	3264	5	US-09-724-676-34558	Sequence 34558, A
c 44	111	9.4	3264	5	US-09-724-676A-34558	Sequence 34558, A
c 45	111	9.4	3280	5	US-09-724-676-34582	Sequence 34582, A

ALIGNMENTS

RESULT 1
US-09-658-659B-7
; Sequence 7, Application US/09658659B
; GENERAL INFORMATION:
; APPLICANT: Stanton, Jr., Vincent P.
; TITLE OF INVENTION: GENE SEQUENCE VARIANCE IN GENES RELATED TO FOLATE METABOLISM HAVING UTILITY IN DETERMINING THE
; TITLE OF INVENTION: TREATMENT OF DISEASE
; FILE REFERENCE: 11926-015001
; CURRENT APPLICATION NUMBER: US/09/658, 659B
; CURRENT FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 09/596, 033
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 09/357, 743
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 09/357, 024
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: 60/093, 484
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1187
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 276, 321, 534, 656
; OTHER INFORMATION: n = c or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 452, 640
; OTHER INFORMATION: n = a or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 492, 625
; OTHER INFORMATION: n = c or a
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 458
; OTHER INFORMATION: nucleotide in position 458 is c, or absent
US-09-658-659B-7

Query Match 100.0%; Score 1187; DB 5; Length 1187;
Best Local Similarity 100.0%; Pred. No. 6.4e-315;
Matches 1187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GATCGGCACTGCACCTCCGCTGGTGAGACGCGAGCTCTGTCTCAAAAAAAAAA 60
|||||

61	AAAAAGACCGCAGGGCTCAACAAACAAAACCTCGAAAAAGCCCTGGCGGTCTTTT	120
138	AAAAAGACCGCAGGGCTCAACAAACAAAACCTCGAAAAAGCCCTGGCGGTCTTTT	197
121	TTTTTTTTTTTTTTTTTTTTTTGGGACAGTCTTGTCTCTGTGCCCAAGGCTGGAGTACAATGGT	180
198	TTTTTTTTTTTTTTTTTTTTTTGGGACAGTCTTGTCTCTGTGCCCAAGGCTGGAGTACAATGGT	257
181	CGGATCTTGGCTCACTCAACCTCTGCTCCAGGTTCAAGCAATTTCTTCTCCCTACAGCC	240
258	CGGATCTTGGCTCACTCAACCTCTGCTCCAGGTTCAAGCAATTTCTTCTCCCTACAGCC	317
241	TCCCAGTAGCACCAACGCCAGCTAAATTTTTTGTANTTTTAGTAGAGACGGGGTTTCAC	300
318	TCCCAGTAGCACCAACGCCAGCTAAATTTTTTGTACTTTTAGTAGAGACGGGGTTTCAC	377
301	CATGTGTCCAGGCTGTCTNGAACTCCTGACCTCAGGTGATCCA CCGGCTTCGGCCGCC	360
378	CATGTGTCCAGGCTGTCTNGAACTCCTGACCTCAGGTGATCCA CCGGCTTCGGCCGCC	437
361	CAAGTACTAGGATTTACAGGCGTAGGCCACCGGCTCAGCGCCCTGGCGGTTTTTAATCA	420
438	CAAGTACTAGGATTTACAGGCGTAGGCCACCGGCTCAGCGCCCTGGCGGTTTTTAATCA	497
421	AGTAGAAAAGCTGCATTAATACACTTGTCTTCNGTTGNTTTCAGTGAAGACGAAGAAATGG	480
498	AGTAGAAAAGCTGCATTAATACACTTGTCTTCNGTTGNTTTCAGTGAAGACGAAGAAATGG	556
481	AAATGCAAATCCTTATTAGTTGTAGGAAAACAGATCTCAACACGACGATTTGTNGACAAG	540
557	AAATGCAAATCCTTATTAGTTGTAGGAAAACAGATCTCAACACGACGATTTGTNGACAAG	616
541	ACGCGAGGAAAACGTGGGAACTGTCTCTGCTGTAGGAAAGGCGCGTGCACGACAGCG	600
617	ACGCGAGGAAAACGTGGGAACTGTCTCTGCTGTAGGAAAGGCGCGTGCACGACAGCG	676
601	TTCCCAAGGGGCGAGTCTTCCGNGCACCGACCTGCTGCTCAGGTTCCCGGTTTCCT	660
677	TTCCCAAGGGGCGAGTCTTCCGNGCACCGACCTGCTGCTCAGGTTCCCGGTTTCCT	736
661	AAGACTCTCAGCTGTGGCCCTGGGCTCGCTGTCTGTCGCACACCGTGGCTCTCGCTTTC	720
737	AAGACTCTCAGCTGTGGCCCTGGGCTCGCTGTCTGTCGCACACCGTGGCTCTCGCTTTC	796
721	CCCTGTGCGCACGCTCTCTAGAGCGGGGCGCCGCGCACCCGCGGCTCTCGCTTTC	780
797	CCCTGTGCGCACGCTCTCTAGAGCGGGGCGCCGCGCACCCGCGGCTCTCGCTTTC	856
781	GAGCGGGGACGCGCCGCGGAAAAGCGCGGGAAGGGTCTTGCACCGCGCCACTTGG	840
857	GAGCGGGGACGCGCCGCGGAAAAGCGCGGGAAGGGTCTTGCACCGCGCCACTTGG	916
841	CCTGCTCTCCGTCGCGCGGACACTTGGGCTGCTCTCCGCGCGGCGCACCTTGCCTG	900
917	CCTGCTCTCCGTCGCGCGGACACTTGGGCTGCTCTCCGCGCGGCGCACCTTGCCTG	976
901	CCTGCTCTCCGTCGCGCGGACACTTGGGCTGCTCTCCGCGCGGCGCACCTTGCCTG	960
977	CCTGCTCTCCGTCGCGCGGACACTTGGGCTGCTCTCCGCGCGGCGCACCTTGCCTG	1036
961	TTGCCCCCGCCGACAGGAGCGGGAACGCCGAGCGCTCTCCGCGCACCGGGAGCTGCAG	1020
1037	TTGCCCCCGCCGACAGGAGCGGGAACGCCGAGCGCTCTCCGCGCACCGGGAGCTGCAG	1096
1021	TACCTGGGGCAGATCCAAACATCTCTCCGCTGCGCGCTCAGAAAGGACGACCGCACGGGC	1080
1097	TACCTGGGGCAGATCCAAACATCTCTCCGCTGCGCGCTCAGAAAGGACGACCGCACGGGC	1156
1081	ACCGGACCCCTGTCTGGTATTTCGGCATGCGAGGCGCGCTACAGCTTGAGAGGTGACCGCG	1140
1157	ACCGGACCCCTGTCTGGTATTTCGGCATGCGAGGCGCGCTACAGCTTGAGAGGTGACCGCG	1216
1141	GGCCCTCTCGGGACGGGTGCGGGGAAAGGAGGAGGCGCGGCTGGGGA	1197

[illegible]

Db 1929 CCACTTGGCTGCTCCGTCCTCCGCGCCACATTGGGCTGCTCCGTCCTCCGCGCCAC 1870

QY 893 TTGCGCTGCTCCGTCCTCCGCGCCACATTGGGCTGCTCCGTCCTCCGCGCCAC 952

Db 1869 TTGCGCTGCTCCGTCCTCCGCGCCACATTGGGCTGCTCCGTCCTCCGCGCCAC 1810

QY 953 GCGCGCTGCTCCGTCCTCCGCGCCACATTGGGCTGCTCCGTCCTCCGCGCCAC 1012

Db 1809 GCGCGCTGCTCCGTCCTCCGCGCCACATTGGGCTGCTCCGTCCTCCGCGCCAC 1750

QY 1013 AGCTGCACTGCTCCGTCCTCCGCGCCACATTGGGCTGCTCCGTCCTCCGCGCCAC 1072

Db 1749 AGCTGCACTGCTCCGTCCTCCGCGCCACATTGGGCTGCTCCGTCCTCCGCGCCAC 1690

QY 1073 GCAAGGGACCGGACCGCTGCTCCGCGCCACATTGGGCTGCTCCGTCCTCCGCGCCAC 1132

Db 1689 GCAAGGGACCGGACCGCTGCTCCGCGCCACATTGGGCTGCTCCGTCCTCCGCGCCAC 1630

QY 1133 A 1133

Db 1629 A 1629

RESULT 5

US-09-724-676-34515/c

Query Match 25.4%; Score 301; DB 5; Length 2042;

Best Local Similarity 100.0%; Pred. No. 3.9e-74;

Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

GENERAL INFORMATION:

APPLICANT: Compugen LTD

TITLE OF INVENTION: Variants of alternative splicing

FILE REFERENCE: 129181.4 Compugen

CURRENT APPLICATION NUMBER: US/09/724,676

CURRENT FILING DATE: 2000-11-28

NUMBER OF SEQ ID NOS: 97222

SOFTWARE: PatentIn version 3.2

SEQ ID NO 34515

LENGTH: 2042

TYPE: DNA

ORGANISM: Homo sapiens

US-09-724-676-34515

Query Match 25.4%; Score 301; DB 5; Length 2042;

Best Local Similarity 100.0%; Pred. No. 3.9e-74;

Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

GENERAL INFORMATION:

APPLICANT: Compugen LTD

TITLE OF INVENTION: Variants of alternative splicing

FILE REFERENCE: 129181.4 Compugen

CURRENT APPLICATION NUMBER: US/09/724,676

CURRENT FILING DATE: 2000-11-28

NUMBER OF SEQ ID NOS: 97222

SOFTWARE: PatentIn version 3.2

SEQ ID NO 34515

LENGTH: 2042

TYPE: DNA

ORGANISM: Homo sapiens

US-09-724-676-34515

QY 833 CCACTTGGCTGCTCCGTCCTCCGCGCCACATTGGGCTGCTCCGTCCTCCGCGCCAC 892

Db 2029 CCACTTGGCTGCTCCGTCCTCCGCGCCACATTGGGCTGCTCCGTCCTCCGCGCCAC 1970

QY 893 TTGCGCTGCTCCGTCCTCCGCGCCACATTGGGCTGCTCCGTCCTCCGCGCCAC 952

Db 1969 TTGCGCTGCTCCGTCCTCCGCGCCACATTGGGCTGCTCCGTCCTCCGCGCCAC 1910

QY 953 GCGCGCTGCTCCGTCCTCCGCGCCACATTGGGCTGCTCCGTCCTCCGCGCCAC 1012

Db 1909 GCGCGCTGCTCCGTCCTCCGCGCCACATTGGGCTGCTCCGTCCTCCGCGCCAC 1850

QY 1013 AGCTGCACTGCTCCGTCCTCCGCGCCACATTGGGCTGCTCCGTCCTCCGCGCCAC 1072

Db 1849 AGCTGCACTGCTCCGTCCTCCGCGCCACATTGGGCTGCTCCGTCCTCCGCGCCAC 1790

QY 1073 GCAAGGGACCGGACCGCTGCTCCGCGCCACATTGGGCTGCTCCGTCCTCCGCGCCAC 1132

Db 1789 GCAAGGGACCGGACCGCTGCTCCGCGCCACATTGGGCTGCTCCGTCCTCCGCGCCAC 1730

QY 1133 A 1133

Db 1729 A 1729

RESULT 6

US-09-724-676-34597/c

Query Match 25.4%; Score 301; DB 5; Length 2391;

Best Local Similarity 100.0%; Pred. No. 3.7e-74;

Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

GENERAL INFORMATION:

APPLICANT: Compugen LTD

TITLE OF INVENTION: Variants of alternative splicing

FILE REFERENCE: 129181.4 Compugen

CURRENT APPLICATION NUMBER: US/09/724,676

CURRENT FILING DATE: 2000-11-28

NUMBER OF SEQ ID NOS: 97222

SOFTWARE: PatentIn version 3.2

SEQ ID NO 34597

LENGTH: 2391

TYPE: DNA

ORGANISM: Homo sapiens

US-09-724-676-34597

QY 833 CCACTTGGCTGCTCCGTCCTCCGCGCCACATTGGGCTGCTCCGTCCTCCGCGCCAC 892

Db 2378 CCACTTGGCTGCTCCGTCCTCCGCGCCACATTGGGCTGCTCCGTCCTCCGCGCCAC 2319

QY 893 TTGCGCTGCTCCGTCCTCCGCGCCACATTGGGCTGCTCCGTCCTCCGCGCCAC 952

Db 2318 TTGCGCTGCTCCGTCCTCCGCGCCACATTGGGCTGCTCCGTCCTCCGCGCCAC 2259

QY 953 GCGCGCTGCTCCGTCCTCCGCGCCACATTGGGCTGCTCCGTCCTCCGCGCCAC 1012

APPLICANT: Compugen LTD

TITLE OF INVENTION: Variants of alternative splicing

FILE REFERENCE: 129181.4 Compugen

CURRENT APPLICATION NUMBER: US/09/724,676A

CURRENT FILING DATE: 2000-11-28

NUMBER OF SEQ ID NOS: 97222

SOFTWARE: PatentIn version 3.2

SEQ ID NO 34515

LENGTH: 2042

TYPE: DNA

ORGANISM: Homo sapiens

US-09-724-676A-34515

Query Match 25.4%; Score 301; DB 5; Length 2042;

Best Local Similarity 100.0%; Pred. No. 3.9e-74;

Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

GENERAL INFORMATION:

APPLICANT: Compugen LTD

TITLE OF INVENTION: Variants of alternative splicing

FILE REFERENCE: 129181.4 Compugen

CURRENT APPLICATION NUMBER: US/09/724,676

CURRENT FILING DATE: 2000-11-28

NUMBER OF SEQ ID NOS: 97222

SOFTWARE: PatentIn version 3.2

SEQ ID NO 34515

LENGTH: 2042

TYPE: DNA

ORGANISM: Homo sapiens

US-09-724-676A-34515

QY 833 CCACTTGGCTGCTCCGTCCTCCGCGCCACATTGGGCTGCTCCGTCCTCCGCGCCAC 892

Db 2029 CCACTTGGCTGCTCCGTCCTCCGCGCCACATTGGGCTGCTCCGTCCTCCGCGCCAC 1970

QY 893 TTGCGCTGCTCCGTCCTCCGCGCCACATTGGGCTGCTCCGTCCTCCGCGCCAC 952

Db 1969 TTGCGCTGCTCCGTCCTCCGCGCCACATTGGGCTGCTCCGTCCTCCGCGCCAC 1910

QY 953 GCGCGCTGCTCCGTCCTCCGCGCCACATTGGGCTGCTCCGTCCTCCGCGCCAC 1012

Db 1909 GCGCGCTGCTCCGTCCTCCGCGCCACATTGGGCTGCTCCGTCCTCCGCGCCAC 1850

QY 1013 AGCTGCACTGCTCCGTCCTCCGCGCCACATTGGGCTGCTCCGTCCTCCGCGCCAC 1072

Db 1849 AGCTGCACTGCTCCGTCCTCCGCGCCACATTGGGCTGCTCCGTCCTCCGCGCCAC 1790

QY 1073 GCAAGGGACCGGACCGCTGCTCCGCGCCACATTGGGCTGCTCCGTCCTCCGCGCCAC 1132

Db 1789 GCAAGGGACCGGACCGCTGCTCCGCGCCACATTGGGCTGCTCCGTCCTCCGCGCCAC 1730

QY 1133 A 1133

Db 1729 A 1729

RESULT 7

US-09-724-676-34597/c

Query Match 25.4%; Score 301; DB 5; Length 2391;

Best Local Similarity 100.0%; Pred. No. 3.7e-74;

Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

GENERAL INFORMATION:

APPLICANT: Compugen LTD

TITLE OF INVENTION: Variants of alternative splicing

FILE REFERENCE: 129181.4 Compugen

CURRENT APPLICATION NUMBER: US/09/724,676

CURRENT FILING DATE: 2000-11-28

NUMBER OF SEQ ID NOS: 97222

SOFTWARE: PatentIn version 3.2

SEQ ID NO 34597

LENGTH: 2391

TYPE: DNA

ORGANISM: Homo sapiens

US-09-724-676-34597

QY 833 CCACTTGGCTGCTCCGTCCTCCGCGCCACATTGGGCTGCTCCGTCCTCCGCGCCAC 892

Db 2378 CCACTTGGCTGCTCCGTCCTCCGCGCCACATTGGGCTGCTCCGTCCTCCGCGCCAC 2319

QY 893 TTGCGCTGCTCCGTCCTCCGCGCCACATTGGGCTGCTCCGTCCTCCGCGCCAC 952

Db 2318 TTGCGCTGCTCCGTCCTCCGCGCCACATTGGGCTGCTCCGTCCTCCGCGCCAC 2259

QY 953 GCGCGCTGCTCCGTCCTCCGCGCCACATTGGGCTGCTCCGTCCTCCGCGCCAC 1012

b 2258 GCCGCCCTTCCCGCCGACAGAGCGGGACGCCGCGCTCCCGCCGACCGGG 2199
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (41)..(41)
; OTHER INFORMATION: n is a,c,g, or t
US-09-724-676-34535

Query Match 25.4%; Score 301; DB 5; Length 2411;
Best Local Similarity 100.0%; Pred. No. 3.7e-74;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1013 AGCTCAGTACTTGGGCGAGATCCACATCTCCGTCGGGCGTCAGAGAGCAGC 1072
b 2198 AGCTCAGTACTTGGGCGAGATCCACATCTCCGTCGGGCGTCAGAGAGCAGC 2139
Y 1073 GCACGGGACCGGACCGCTTGGGATTCGGGATCGAGCGCGCTACAGCTGAGAGGTG 1132
b 2138 GCACGGGACCGGACCGCTTGGGATTCGGGATCGAGCGCGCTACAGCTGAGAGGTG 2079
Y 1133 A 1133
b 2078 A 2078

RESULT 8
US-09-724-676A-34597/c
Sequence 34597, Application US/09724676A
GENERAL INFORMATION:
APPLICANT: Compugen LTD
TITLE OF INVENTION: Variants of alternative splicing
FILE REFERENCE: 129181.4 Compugen
CURRENT APPLICATION NUMBER: US/09/724,676A
CURRENT FILING DATE: 2000-11-28
NUMBER OF SEQ ID NOS: 97222
SOFTWARE: PatentIn version 3.2
SEQ ID NO 34597
LENGTH: 2391
TYPE: DNA
ORGANISM: Homo sapiens
US-09-724-676A-34597

Query Match 25.4%; Score 301; DB 5; Length 2391;
Best Local Similarity 100.0%; Pred. No. 3.7e-74;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 833 CCACCTTGGGCTGCTCCGTCGCCGCGGACATTCGCTCCGTCGCCGCGCCAC 892
b 2378 CCACCTTGGGCTGCTCCGTCGCCGCGGACATTCGCTCCGTCGCCGCGCCAC 2319
Y 893 TTGCGCTGCTCCGTCGCCGCGGACATTCGCTCCGTCGCCGCGCCAC 952
b 2318 TTGCGCTGCTCCGTCGCCGCGGACATTCGCTCCGTCGCCGCGCCAC 2259
Y 953 GCGGCGCTTGGCCCGCGGACATTCGCTCCGTCGCCGCGCCAC 1012
b 2258 GCGGCGCTTGGCCCGCGGACATTCGCTCCGTCGCCGCGCCAC 2199
Y 1013 AGCTCAGTACTTGGGCGAGATCCACATCTCCGTCGGGCGTCAGAGAGCAGC 1072
b 2198 AGCTCAGTACTTGGGCGAGATCCACATCTCCGTCGGGCGTCAGAGAGCAGC 2139
Y 1073 GCACGGGACCGGACCGCTTGGGATTCGGGATCGAGCGCGCTACAGCTGAGAGGTG 1132
b 2138 GCACGGGACCGGACCGCTTGGGATTCGGGATCGAGCGCGCTACAGCTGAGAGGTG 2079
Y 1133 A 1133
b 2078 A 2078

RESULT 9
US-09-724-676-34535/c
Sequence 34535, Application US/09724676
GENERAL INFORMATION:
APPLICANT: Compugen LTD
TITLE OF INVENTION: Variants of alternative splicing
FILE REFERENCE: 129181.4 Compugen
CURRENT APPLICATION NUMBER: US/09/724,676
CURRENT FILING DATE: 2000-11-28
NUMBER OF SEQ ID NOS: 97222
SOFTWARE: PatentIn version 3.2
SEQ ID NO 34535

Query Match 25.4%; Score 301; DB 5; Length 2411;
Best Local Similarity 100.0%; Pred. No. 3.7e-74;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 833 CCACCTTGGGCTGCTCCGTCGCCGCGGACATTCGCTCCGTCGCCGCGCCAC 892
b 2398 CCACCTTGGGCTGCTCCGTCGCCGCGGACATTCGCTCCGTCGCCGCGCCAC 2339
Y 893 TTGCGCTGCTCCGTCGCCGCGGACATTCGCTCCGTCGCCGCGCCAC 952
b 2338 TTGCGCTGCTCCGTCGCCGCGGACATTCGCTCCGTCGCCGCGCCAC 2279
Y 953 GCGGCGCTTGGCCCGCGGACATTCGCTCCGTCGCCGCGCCAC 1012
b 2158 GCGGCGCTTGGCCCGCGGACATTCGCTCCGTCGCCGCGCCAC 2159
Y 1133 A 1133
b 2098 A 2098

LENGTH: 2411
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (41)..(41)
OTHER INFORMATION: n is a,c,g, or t
US-09-724-676-34535

Query Match 25.4%; Score 301; DB 5; Length 2411;
Best Local Similarity 100.0%; Pred. No. 3.7e-74;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 833 CCACCTTGGGCTGCTCCGTCGCCGCGGACATTCGCTCCGTCGCCGCGCCAC 892
b 2398 CCACCTTGGGCTGCTCCGTCGCCGCGGACATTCGCTCCGTCGCCGCGCCAC 2339
Y 893 TTGCGCTGCTCCGTCGCCGCGGACATTCGCTCCGTCGCCGCGCCAC 952
b 2338 TTGCGCTGCTCCGTCGCCGCGGACATTCGCTCCGTCGCCGCGCCAC 2279
Y 953 GCGGCGCTTGGCCCGCGGACATTCGCTCCGTCGCCGCGCCAC 1012
b 2278 GCGGCGCTTGGCCCGCGGACATTCGCTCCGTCGCCGCGCCAC 2219
Y 1013 AGCTCAGTACTTGGGCGAGATCCACATCTCCGTCGGGCGTCAGAGAGCAGC 1072
b 2218 AGCTCAGTACTTGGGCGAGATCCACATCTCCGTCGGGCGTCAGAGAGCAGC 2159
Y 1073 GCACGGGACCGGACCGCTTGGGATTCGGGATCGAGCGCGCTACAGCTGAGAGGTG 1132
b 2158 GCACGGGACCGGACCGCTTGGGATTCGGGATCGAGCGCGCTACAGCTGAGAGGTG 2099
Y 1133 A 1133
b 2098 A 2098

RESULT 10
US-09-724-676A-34535/c
Sequence 34535, Application US/09724676A
GENERAL INFORMATION:
APPLICANT: Compugen LTD
TITLE OF INVENTION: Variants of alternative splicing
FILE REFERENCE: 129181.4 Compugen
CURRENT APPLICATION NUMBER: US/09/724,676A
CURRENT FILING DATE: 2000-11-28
NUMBER OF SEQ ID NOS: 97222
SOFTWARE: PatentIn version 3.2
SEQ ID NO 34535
LENGTH: 2411
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (41)..(41)
OTHER INFORMATION: n is a,c,g, or t
US-09-724-676A-34535

Query Match 25.4%; Score 301; DB 5; Length 2411;
Best Local Similarity 100.0%; Pred. No. 3.7e-74;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 833 CCACCTTGGGCTGCTCCGTCGCCGCGGACATTCGCTCCGTCGCCGCGCCAC 892
b 2398 CCACCTTGGGCTGCTCCGTCGCCGCGGACATTCGCTCCGTCGCCGCGCCAC 2339
Y 893 TTGCGCTGCTCCGTCGCCGCGGACATTCGCTCCGTCGCCGCGCCAC 952
b 2338 TTGCGCTGCTCCGTCGCCGCGGACATTCGCTCCGTCGCCGCGCCAC 2279
Y 953 GCGGCGCTTGGCCCGCGGACATTCGCTCCGTCGCCGCGCCAC 1012
b 2158 GCGGCGCTTGGCCCGCGGACATTCGCTCCGTCGCCGCGCCAC 2159
Y 1133 A 1133
b 2098 A 2098

Db	2278	GCCGGCCCTTGGCCCCCGCCGACAGAGCGGGACGCCGCGTCCGCGCGACCGGG	2219
Qy	1013	AGCTGACGTACTTGGGGCAGATCCAAACATCTCCGCTGCGGCGTCAAGGAAGACGACC	1072
Db	2218	AGCTGACGTACTTGGGGCAGATCCAAACATCTCCGCTGCGGCGTCAAGGAAGACGACC	2159
Qy	1073	GCAAGGACACCGCACCTCTGCGTATTCGGCATGCAAGGCGCTACAGCTGAGAGGTG	1132
Db	2158	GCAAGGACACCGCACCTCTGCGTATTCGGCATGCAAGGCGCTACAGCTGAGAGGTG	2099
Qy	1133	A	1133
Db	2098	A	2098
RESULT 11			
US-09-724-676-34589/c			
; Sequence 34589, Application US/09724676			
; GENERAL INFORMATION:			
; APPLICANT: Compugen LTD			
; TITLE OF INVENTION: Variants of alternative splicing			
; FILE REFERENCE: 129181.4 Compugen			
; CURRENT APPLICATION NUMBER: US/09/724,676			
; CURRENT FILING DATE: 2000-11-28			
; NUMBER OF SEQ ID NOS: 97222			
; SOFTWARE: PatentIn version 3.2			
; SEQ ID NO 34589			
; LENGTH: 2491			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-09-724-676-34589			
Query Match			
Best Local Similarity 25.4%; Score 301; DB 5; Length 2491;			
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	833	CCACTTGGCTTGCCTCCGTCCTCCGCGCCACTTGGGCTGCTCCGTCGGCGCGCCAC	892
Db	2478	CCACTTGGCTTGCCTCCGTCCTCCGCGCCACTTGGGCTGCTCCGTCGGCGCGCCAC	2419
Qy	893	TTGCGCTTGCCTCCGTCCTCCGCGCCACTTGGGCTGCTCCGTCGGCGCGCCAC	952
Db	2418	TTGCGCTTGCCTCCGTCCTCCGCGCCACTTGGGCTGCTCCGTCGGCGCGCCAC	2359
Qy	953	GCGGCGCTTGCCTCCGTCCTCCGCGCCACTTGGGCTGCTCCGTCGGCGCGCCAC	1012
Db	2358	GCGGCGCTTGCCTCCGTCCTCCGCGCCACTTGGGCTGCTCCGTCGGCGCGCCAC	2299
Qy	1013	AGCTGACGTACTTGGGGCAGATCCAAACATCTCCGCTGCGGCGTCAGGAAGACGACC	1072
Db	2298	AGCTGACGTACTTGGGGCAGATCCAAACATCTCCGCTGCGGCGTCAGGAAGACGACC	2239
Qy	1073	GCAAGGACACCGCACCTCTGCGTATTCGGCATGCAAGGCGCTACAGCTGAGAGGTG	1132
Db	2238	GCAAGGACACCGCACCTCTGCGTATTCGGCATGCAAGGCGCTACAGCTGAGAGGTG	2179
Qy	1133	A	1133
Db	2178	A	2178
RESULT 13			
US-09-724-676-34548/c			
; Sequence 34548, Application US/09724676			
; GENERAL INFORMATION:			
; APPLICANT: Compugen LTD			
; TITLE OF INVENTION: Variants of alternative splicing			
; FILE REFERENCE: 129181.4 Compugen			
; CURRENT APPLICATION NUMBER: US/09/724,676			
; CURRENT FILING DATE: 2000-11-28			
; NUMBER OF SEQ ID NOS: 97222			
; SOFTWARE: PatentIn version 3.2			
; SEQ ID NO 34548			
; LENGTH: 2527			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; NAME/KEY: misc feature			
; LOCATION: (41)..(41)			
; OTHER INFORMATION: n is a,c,g, or t			
US-09-724-676-34548			
Query Match			
Best Local Similarity 25.4%; Score 301; DB 5; Length 2527;			
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	833	CCACTTGGCTTGCCTCCGTCCTCCGCGCGCCACTTGGGCTGCTCCGTCGGCGCGCCAC	892
Db	2514	CCACTTGGCTTGCCTCCGTCCTCCGCGCGCCACTTGGGCTGCTCCGTCGGCGCGCCAC	2455
Qy	893	TTGCGCTTGCCTCCGTCCTCCGCGCGCCACTTGGGCTGCTCCGTCGGCGCGCCAC	952
Db	2454	TTGCGCTTGCCTCCGTCCTCCGCGCGCCACTTGGGCTGCTCCGTCGGCGCGCCAC	2395
Qy	953	GCGGCGCTTGCCTCCGTCCTCCGCGCGCCACTTGGGCTGCTCCGTCGGCGCGCCAC	1012
Db	2394	GCGGCGCTTGCCTCCGTCCTCCGCGCGCCACTTGGGCTGCTCCGTCGGCGCGCCAC	2335
Qy	1013	AGCTGACGTACTTGGGGCAGATCCAAACATCTCCGCTGCGGCGTCAGGAAGACGACC	1072

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 29, 2002, 00:35:25 ; Search time 2462.49 Seconds
(without alignments)
12119.489 Million cell updates/sec

Title: US-09-963-333-7

Perfect score: 1187
Sequence: 1 gatcgccactgactcca.....gagggagcgcgctgggga 1187

Scoring table:

OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched:

24791104 seqs, 12571243825 residues

Word size : 15

Total number of hits satisfying chosen parameters: 2685365

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Pending Patents_NA_Main:*

- 1: /cgn2_6/ptodata/2/pna/pctus COMB.seq:*
- 2: /cgn2_6/ptodata/2/pna/US06 COMB.seq:*
- 3: /cgn2_6/ptodata/2/pna/US07 COMB.seq:*
- 4: /cgn2_6/ptodata/2/pna/US08 COMB.seq:*
- 5: /cgn2_6/ptodata/2/pna/US09 COMB.seq:*
- 6: /cgn2_6/ptodata/2/pna/US08 COMB.seq:*
- 7: /cgn2_6/ptodata/2/pna/US08 COMB.seq:*
- 8: /cgn2_6/ptodata/2/pna/US08 COMB.seq:*
- 9: /cgn2_6/ptodata/2/pna/US08 COMB.seq:*
- 10: /cgn2_6/ptodata/2/pna/US08 COMB.seq:*
- 11: /cgn2_6/ptodata/2/pna/US08 COMB.seq:*
- 12: /cgn2_6/ptodata/2/pna/US08 COMB.seq:*
- 13: /cgn2_6/ptodata/2/pna/US08 COMB.seq:*
- 14: /cgn2_6/ptodata/2/pna/US09 COMB.seq:*
- 15: /cgn2_6/ptodata/2/pna/US09 COMB.seq:*
- 16: /cgn2_6/ptodata/2/pna/US09 COMB.seq:*
- 17: /cgn2_6/ptodata/2/pna/US09 COMB.seq:*
- 18: /cgn2_6/ptodata/2/pna/US09 COMB.seq:*
- 19: /cgn2_6/ptodata/2/pna/US09 COMB.seq:*
- 20: /cgn2_6/ptodata/2/pna/US09 COMB.seq:*
- 21: /cgn2_6/ptodata/2/pna/US09 COMB.seq:*
- 22: /cgn2_6/ptodata/2/pna/US09 COMB.seq:*
- 23: /cgn2_6/ptodata/2/pna/US09 COMB.seq:*
- 24: /cgn2_6/ptodata/2/pna/US09 COMB.seq:*
- 25: /cgn2_6/ptodata/2/pna/US09 COMB.seq:*
- 26: /cgn2_6/ptodata/2/pna/US09 COMB.seq:*
- 27: /cgn2_6/ptodata/2/pna/US09 COMB.seq:*
- 28: /cgn2_6/ptodata/2/pna/US09 COMB.seq:*
- 29: /cgn2_6/ptodata/2/pna/US09 COMB.seq:*
- 30: /cgn2_6/ptodata/2/pna/US09 COMB.seq:*
- 31: /cgn2_6/ptodata/2/pna/US09 COMB.seq:*
- 32: /cgn2_6/ptodata/2/pna/US09 COMB.seq:*
- 33: /cgn2_6/ptodata/2/pna/US09 COMB.seq:*
- 34: /cgn2_6/ptodata/2/pna/US09 COMB.seq:*
- 35: /cgn2_6/ptodata/2/pna/US09 COMB.seq:*
- 36: /cgn2_6/ptodata/2/pna/US09 COMB.seq:*
- 37: /cgn2_6/ptodata/2/pna/US09 COMB.seq:*
- 38: /cgn2_6/ptodata/2/pna/US09 COMB.seq:*
- 39: /cgn2_6/ptodata/2/pna/US09 COMB.seq:*
- 40: /cgn2_6/ptodata/2/pna/US09 COMB.seq:*
- 41: /cgn2_6/ptodata/2/pna/US09 COMB.seq:*
- 42: /cgn2_6/ptodata/2/pna/US09 COMB.seq:*
- 43: /cgn2_6/ptodata/2/pna/US09 COMB.seq:*

- 44: /cgn2_6/ptodata/2/pna/US6000 COMB.seq:*
- 45: /cgn2_6/ptodata/2/pna/US6001 COMB.seq:*
- 46: /cgn2_6/ptodata/2/pna/US6002 COMB.seq:*
- 47: /cgn2_6/ptodata/2/pna/US6003 COMB.seq:*
- 48: /cgn2_6/ptodata/2/pna/US6004 COMB.seq:*
- 49: /cgn2_6/ptodata/2/pna/US6005 COMB.seq:*
- 50: /cgn2_6/ptodata/2/pna/US6006 COMB.seq:*
- 51: /cgn2_6/ptodata/2/pna/US6007 COMB.seq:*
- 52: /cgn2_6/ptodata/2/pna/US6008 COMB.seq:*
- 53: /cgn2_6/ptodata/2/pna/US6009 COMB.seq:*
- 54: /cgn2_6/ptodata/2/pna/US6010 COMB.seq:*
- 55: /cgn2_6/ptodata/2/pna/US6011 COMB.seq:*
- 56: /cgn2_6/ptodata/2/pna/US6012 COMB.seq:*
- 57: /cgn2_6/ptodata/2/pna/US6013 COMB.seq:*
- 58: /cgn2_6/ptodata/2/pna/US6014 COMB.seq:*
- 59: /cgn2_6/ptodata/2/pna/US6015 COMB.seq:*
- 60: /cgn2_6/ptodata/2/pna/US6016 COMB.seq:*
- 61: /cgn2_6/ptodata/2/pna/US6017 COMB.seq:*
- 62: /cgn2_6/ptodata/2/pna/US6018 COMB.seq:*
- 63: /cgn2_6/ptodata/2/pna/US6019 COMB.seq:*
- 64: /cgn2_6/ptodata/2/pna/US6020 COMB.seq:*
- 65: /cgn2_6/ptodata/2/pna/US6021 COMB.seq:*
- 66: /cgn2_6/ptodata/2/pna/US6022 COMB.seq:*
- 67: /cgn2_6/ptodata/2/pna/US6023 COMB.seq:*
- 68: /cgn2_6/ptodata/2/pna/US6024 COMB.seq:*
- 69: /cgn2_6/ptodata/2/pna/US6025 COMB.seq:*
- 70: /cgn2_6/ptodata/2/pna/US6026 COMB.seq:*
- 71: /cgn2_6/ptodata/2/pna/US6027 COMB.seq:*
- 72: /cgn2_6/ptodata/2/pna/US6028 COMB.seq:*
- 73: /cgn2_6/ptodata/2/pna/US6029 COMB.seq:*
- 74: /cgn2_6/ptodata/2/pna/US6030 COMB.seq:*
- 75: /cgn2_6/ptodata/2/pna/US6031 COMB.seq:*
- 76: /cgn2_6/ptodata/2/pna/US6032 COMB.seq:*
- 77: /cgn2_6/ptodata/2/pna/US6033 COMB.seq:*
- 78: /cgn2_6/ptodata/2/pna/US6034 COMB.seq:*
- 79: /cgn2_6/ptodata/2/pna/US6035 COMB.seq:*
- 80: /cgn2_6/ptodata/2/pna/US6036 COMB.seq:*
- 81: /cgn2_6/ptodata/2/pna/US6037 COMB.seq:*
- 82: /cgn2_6/ptodata/2/pna/US6038 COMB.seq:*
- 83: /cgn2_6/ptodata/2/pna/US6039 COMB.seq:*
- 84: /cgn2_6/ptodata/2/pna/US6040 COMB.seq:*
- 85: /cgn2_6/ptodata/2/pna/US6041 COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1187	100.0	1187	25	US-09-658-659-7 Sequence 7, Appli
2	1187	100.0	1187	36	US-09-962-665-7 Sequence 7, Appli
3	1187	100.0	1187	36	US-09-962-677-7 Sequence 7, Appli
4	1187	100.0	1187	36	US-09-963-333-7 Sequence 7, Appli
5	811	68.3	18597	25	US-09-658-659-8 Sequence 8, Appli
6	811	68.3	18597	36	US-09-962-665-8 Sequence 8, Appli
7	811	68.3	18597	36	US-09-962-677-8 Sequence 8, Appli
8	811	68.3	18597	36	US-09-963-333-8 Sequence 8, Appli
9	658	55.4	18596	22	US-09-577-266-11 Sequence 11, Appl
10	658	55.4	18596	24	US-09-631-275-140 Sequence 140, App
11	658	55.4	18596	33	US-09-880-107-1590 Sequence 1590, Ap
12	658	55.4	18596	36	US-09-934-531-124 Sequence 124, App
13	658	55.4	18596	36	US-09-954-531-348 Sequence 348, App
14	658	55.4	18596	36	US-09-967-768A-119 Sequence 119, App
15	297	25.0	1536	1	PCT-US02-18947-556 Sequence 556, App
16	297	25.0	1536	1	PCT-US99-12080-1 Sequence 792, App
17	297	25.0	1536	18	US-09-442-384A-792 Sequence 792, App
18	297	25.0	1536	18	US-09-442-384B-732 Sequence 141, App
19	297	25.0	1536	24	US-09-631-275-141 Sequence 6, Appli
20	297	25.0	1536	25	US-09-658-659-6 Sequence 11, Appli
21	297	25.0	1536	33	US-09-877-178-11 Sequence 11, Appli

LOCATION: 14547
OTHER INFORMATION: nucleotide in position 14547 is t, or absent
US-09-658-659-8

Query Match 68.3%; Score 811; DB 25; Length 18597;
Best Local Similarity 99.5%; Pred. No. 8.8e-93;
Matches 1181; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 1 GATCGCGCCACTGCACTCCAGCCTGGGTGAGAGAGAGAGACTCTGTCTCAAAAAA 60
DB 78 GATCGCGCCACTGCACTCCAGCCTGGGTGAGAGAGAGAGACTCTGTCTCAAAAAA 137
QY 61 AAAAAACCGCCAGCGCTCAAAACAAAAACCTCGGAAAGCCCTGGCGGCTCTTTT 120
DB 138 AAAAAACCGCCAGCGCTCAAAACAAAAACCTCGGAAAGCCCTGGCGGCTCTTTT 197
QY 121 TTTTCTTTTCTTTTCTTTTGGGACAGTCTTGTCTGTCTGCGCCAGGCTGGAGTCAATGT 180
DB 198 TTTTCTTTTCTTTTCTTTTGGGACAGTCTTGTCTGTCTGCGCCAGGCTGGAGTCAATGT 257
QY 181 CGGATCTGGCTCACTGCAACTCTGCTCCAGGTTCAAGCAATCTTCTGCTCAGCC 240
DB 258 CGGATCTGGCTCACTGCAACTCTGCTCCAGGTTCAAGCAATCTTCTGCTCAGCC 317
QY 241 TCCCAAGTAGCACCACCGCCAGCTAAATTTTGTANTTTTGTAGTAGACGGGGTTTCA 300
DB 318 TCCCAAGTAGCACCACCGCCAGCTAAATTTTGTACTTTTGTAGTAGACGGGGTTTCA 377
QY 301 CATGTGTCCAGGCTGCTTGAACCTCTGACTCAGGTTGATCCACCGGCTCGGCCCC 360
DB 378 CATGTGTCCAGGCTGCTTGAACCTCTGACTCAGGTTGATCCACCGGCTCGGCCCC 437
QY 361 CAAAGTACTAGGATTACAGGCTGAGCCACCGGCTCCAGCGCCCTGGCGGTTTAA 420
DB 438 CAAAGTACTAGGATTACAGGCTGAGCCACCGGCTCCAGCGCCCTGGCGGTTTAA 497
QY 421 AGTAGAAAAGCTGCATTAACACCTGCTTCGCTTCGCTTCGCTTCGCTTCGCTTC 480
DB 498 AGTAGAAAAGCTGCATTAACACCTGCTTCGCTTCGCTTCGCTTCGCTTCGCTTC 556
QY 481 AATGCAATATCTTATAGTTAGTAAACAGATCTCAACAGCAGGTTTGTGACAG 540
DB 557 AATGCAATATCTTATAGTTAGTAAACAGATCTCAACAGCAGGTTTGTGACAG 616
QY 541 ACCGAGGAAAAGCTGGGAACTGTGCTGGCTTAGAGAGCGCGCTCGACACAGCG 600
DB 617 ACCGAGGAAAAGCTGGGAACTGTGCTGGCTTAGAGAGCGCGCTCGACACAGCG 676
QY 601 TTCCCAAGGCGCAGTCTCTCCNGCCACCGCACTGCTCCAGGTTCCCGGGTTCCT 660
DB 677 TTCCCAAGGCGCAGTCTCTCTCCNGCCACCGCACTGCTCCAGGTTCCCGGGTTCCT 736
QY 661 AAGACTCTAGCTGTGGCCCTGGCTCGTTCGCTTCGCTTCGCTTCGCTTCGCTTC 720
DB 737 AAGACTCTAGCTGTGGCCCTGGCTCGTTCGCTTCGCTTCGCTTCGCTTCGCTTC 796
QY 721 CCGTGGCGCAGCTCTTAGAGCGGGGGCGCGGACCGCCGCGAGCAGGAAGAGCG 780
DB 797 CCGTGGCGCAGCTCTTAGAGCGGGGGCGCGGACCGCCGCGAGCAGGAAGAGCG 856
QY 781 GAGCGGGGAGCGCGCGGAAAAGCGCGGGAAGGGGTCCTGACCGCGCCACTTGG 840
DB 857 GAGCGGGGAGCGCGCGGAAAAGCGCGGGAAGGGGTCCTGACCGCGCCACTTGG 916
QY 841 CCGCTCCGTCGCGCGCGCCACTTGGCTTCGCTTCGCTTCGCTTCGCTTCGCTTC 900
DB 917 CCGCTCCGTCGCGCGCGCCACTTGGCTTCGCTTCGCTTCGCTTCGCTTCGCTTC 976
QY 901 CCGCTCCGTCGCGCGCGCCACTTGGCTTCGCTTCGCTTCGCTTCGCTTCGCTTC 960
DB 977 CCGCTCCGTCGCGCGCGCCACTTGGCTTCGCTTCGCTTCGCTTCGCTTCGCTTC 1036
QY 961 TTGCCCCCGCCGACAGAGCGGGAAGCGCGCGGCTCCGCGGACGCGGAGCTGCAG 1020

DB 1037 TTGCCCCCGCCGACAGAGCGGGAAGCGCGGCTCCGCGGACGCGGAGCTGCAG 1096
QY 1021 TACTCGGCGAGATCCAAACATCTTCGCTCGCGCTGCGGCGTCAAGAAAGGACGACCGAGCGGC 1080
DB 1097 TACTCGGCGAGATCCAAACATCTTCGCTCGCGCTGCGGCGTCAAGAAAGGACGACCGAGCGGC 1156
QY 1081 ACCGCGACCCCTGCTCGGCTTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTC 1140
DB 1157 ACCGCGACCCCTGCTCGGCTTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTC 1216
QY 1141 GCGCCCTGCGGACGCGGCTGCGGGAAGGAGGAGGCGGCGGCTTGAGGA 1187
DB 1217 GCGCCCTGCGGACGCGGCTGCGGGAAGGAGGAGGCGGCGGCTTGAGGA 1263

RESULT 6
US-09-962-665-8
; Sequence 8, Application US/09962665
; GENERAL INFORMATION:
; APPLICANT: Stanton, Jr., Vincent P.
; TITLE OF INVENTION: POLY(POLYGLUTAMATE SYNTHETASE GENE SEQUENCE
; TITLE OF INVENTION: VARIANCES HAVING UTILITY IN DETERMINING THE
; TITLE OF INVENTION: TREATMENT OF DISEASE
; FILE REFERENCE: 11926-015004
; CURRENT APPLICATION NUMBER: US/09/962,665
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/658,659
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 09/596,033
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 09/357,743
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 09/357,024
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: 60/093,484
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 18597
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 701, 13751
; OTHER INFORMATION: n = c or a
; NAME/KEY: misc_feature
; LOCATION: 716, 1293, 2401, 2429, 2618, 3083, 3125, 3635, 4256, 4898,
; LOCATION: 5062, 5167, 11069, 13298, 14479, 14730, 14796, 15344, 15450,
; LOCATION: 15503, 15590, 15840, 16149
; OTHER INFORMATION: n = a or g
; NAME/KEY: misc_feature
; LOCATION: 732, 1379, 1590, 2488, 3212, 5006, 11238, 11422, 11686,
; LOCATION: 12598, 13171, 13645, 13782, 13806, 13813, 14586, 14788,
; LOCATION: 15042, 15546, 15770
; OTHER INFORMATION: n = c or t
; NAME/KEY: misc_feature
; LOCATION: 1322, 1688
; OTHER INFORMATION: n = c or g
; NAME/KEY: misc_feature
; LOCATION: 2594, 11293, 16199, 16203
; OTHER INFORMATION: n = g or t
; NAME/KEY: misc_feature
; LOCATION: 3619
; OTHER INFORMATION: n = a or t
; NAME/KEY: misc_feature
; LOCATION: 14547
; OTHER INFORMATION: nucleotide in position 14547 is t, or absent
US-09-962-665-8

Query Match 68.3%; Score 811; DB 36; Length 18597;
Best Local Similarity 99.5%; Pred. No. 8.8e-93;

	Matches 1181;	Conservative	0;	Mismatches	5;	Indels	1;	Gaps	1;
2Y	1	GATCGCGCACTGCACCTCCAGCCTCGGCTGAGAGCGAGACTCTGTCTCAAAAAA	60						
Db	78	GATCGCGCACTGCACCTCCAGCCTCGGCTGAGAGCGAGACTCTGTCTCAAAAAA	137						
2Y	61	AAAAAGCCGCGAGGCTCAACAAAAA	120						
Db	138	AAAAAGACCGCCAGGCTCAACAAAAA	197						
2Y	121	TTTTTTTTTTTTTTTTTTTGGACAGTCTTCTCTGTGCGCCAGGCTGGAGTACAATG	180						
Db	198	TTTTTTTTTTTTTTTTTTTGGGACAGTCTTCTCTGTGCGCCAGGCTGGAGTACAATG	257						
QY	181	CGGAFCTTGGCTCACTGCAACCTCTGCCCTCCAGGTTCAAGCAATCTTCTGCTCAGCC	240						
Db	258	CGGATCTTGGCTCACTGCAACCTCTGCCCTCCAGGTTCAAGCAATCTTCTGCTCAGCC	317						
QY	241	TCCGAATGAGCAACGCGGAGCTAATTTTGTGTGANTTTAGTAGAGCGGGGTTTCA	300						
Db	318	TCCGAATGAGCAACGCGGAGCTAATTTTGTGTGANTTTAGTAGAGCGGGGTTTCA	377						
QY	301	CAATGTTGTCAGGCTGCTGACAACTTCTGACCTCAGGTGATCAACCGCTCGGCCCC	360						
Db	378	CAATGTTGTCAGGCTGCTGACAACTTCTGACCTCAGGTGATCAACCGCTCGGCCCC	437						
QY	361	CAAAGTACTAGGATTAACGAGCGTGAGCCACCGCTCCAGCGCTTGGCGGTTTTTAATCA	420						
Db	438	CAAAGTACTAGGATTAACGAGCGTGAGCCACCGCTCCAGCGCTTGGCGGTTTTTAATCA	497						
QY	421	AGTAGAAGAGCTGCATTATACAATTGCTTCTGTTGNTTCTGAGTAGAAGAAATGG	480						
Db	498	AGTAGAAGAGCTGCATTATACCATTGCTTCGGTTGC-TTCAGTGAAGCAAGAAATGG	556						
QY	481	AAATGCAAAATCCTTATATAGTTGTAGGAAACAGATCTCAAAACAGCAGTTTGTGACAAG	540						
Db	557	AAATGCAAAATCCTTATATAGTTGTAGGAAACAGATCTCAAAACAGCAGTTTGTGACAAG	616						
QY	541	ACCGCAGGAAAACTGGGAACTGTGCTGCTGGCTTAGAGAGCGCGGTGCAACAGACGG	600						
Db	617	ACCGCAGGAAAACTGGGAACTGTGCTGCTGGCTTAGAGAGCGCGGTGCAACAGACGG	676						
QY	601	TTCCGAAGAGGCGAGTCTTCCGNGCCAACGACACTGNTCCAGGTTCCGCGGTTTCT	660						
Db	677	TTCCGAAGAGGCGAGTCTTCCGNGCCAACGACACTGNTCCAGGTTCCGCGGTTTCT	736						
QY	661	AAGACTCTCAGCTGTGGCCCTGGGCTCGTCTGTGCGCACACCGTGGCTCTCGGTTTC	720						
Db	737	AAGACTCTCAGCTGTGGCCCTGGGCTCGTCTGTGCGCACACCGTGGCTCTCGGTTTC	796						
QY	721	CCCTGGCGCACTCTCTTAGAGCGGGGCGCGCGCACCCCGCGCAGCAGGAGAGCGG	780						
Db	797	CCCTGGCGCACTCTCTTAGAGCGGGGCGCGCGCACCCCGCGCAGCAGGAGAGCGG	856						
QY	781	GAGCGCGGAGCGCGCGCGGAAAAAGCGCGGAAAGGGTCTGTGCCACCGCGCACTTGG	840						
Db	857	GAGCGCGGAGCGCGCGCGGAAAAAGCGCGGAAAGGGTCTGTGCCACCGCGCACTTGG	916						
QY	841	CTTGCTCTCGGTCGCGCGCGGCACTTGGCTGTGCTCGTCCGCGCGGCACTTGCCTG	900						
Db	917	CTTGCTCTCGGTCGCGCGCGGCACTTGGCTGTGCTCGTCCGCGCGGCACTTGCCTG	976						
QY	901	CTTGCTCTCGGTCGCGCGGCACTTGGCTGTGCGGCTGTGGAGCTTGGCGCGCGGCC	960						
Db	977	CTTGCTCTCGGTCGCGCGGCACTTGGCTGTGGCGGCTTGGAGCTTGGCGCGCGGCC	1036						
QY	961	TTGCCCCCGCGCGCACAGGAGCGGAGCCCGAGCCGCGCTCCGCGCACGGGAGCTGCAG	1020						
Db	1037	TTGCCCCCGCGCGCACAGGAGCGGAGCCCGAGCCGCGCTCCGCGCACGGGAGCTGCAG	1096						
QY	1021	TACCTGGGCGAGTCCACACATCTCGCTGCGGCTCAAGAAAGGACGACGACGCGG	1080						
Db	1097	TACCTGGGCGAGTCCACACATCTCGCTGCGGCTCAAGAAAGGACGACGACGCGG	1156						

Qy	1081	ACCGGCAACCCCTGCTGGTATTTCCGATATGCAAGCGCGCTACAGCCTGAGAGGTGACGCGCG	1140
Db	1157	ACCGGCAACCCCTGTGCGTATTTCGATATGCGATGACGCGGCTTACGCTTGAGAGGTGACGCGCG	1216
Qy	1141	GGCCCCCTGCGGGAACGGTTCGCGGAAGGAGGAGCGCGCTGGGGA	1187
Db	1217	GGCCCCCTGCGGGAACGGTTCGCGGAAGGAGGAGCGCGCTGGGGA	1263

```

RESULT 7
US-09-962-677-8
; Sequence 8, Application US/09962677
; GENERAL INFORMATION:
; APPLICANT: Stanton, Jr., Vincent P.
; TITLE OF INVENTION: THYMIDINE PHOSPHORYLASE GENE SEQUENCE
; TITLE OF INVENTION: VARIANCES HAVING UTILITY IN DETERMINING
; TITLE OF INVENTION: THE TREATMENT OF DISEASE
; FILE REFERENCE: 11926-015003
; CURRENT APPLICATION NUMBER: US/09/962,677
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/658,659
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 09/596,033
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 09/357,743
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 09/357,024
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: 60/093,484
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 18597
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 701..13751
; OTHER INFORMATION: n = c or a
; NAME/KEY: misc.feature
; LOCATION: 716..1293, 2401..2429, 2618, 3083, 3125, 3635, 4256, 4898,
; LOCATION: 5062, 5167, 11069, 13298, 14479, 14730, 14795, 15344, 15450,
; LOCATION: 15503, 15590, 15840, 16149
; OTHER INFORMATION: n = a or g
; NAME/KEY: misc.feature
; LOCATION: 732, 1379, 1590, 2488, 3212, 5006, 11238, 11422, 11686,
; LOCATION: 12598, 13171, 13645, 13782, 13806, 13813, 14586, 14788,
; LOCATION: 15042, 15546, 15770
; OTHER INFORMATION: n = c or t
; NAME/KEY: misc.feature
; LOCATION: 1322, 1688
; OTHER INFORMATION: n = c or g
; NAME/KEY: misc.feature
; LOCATION: 2594, 11293, 16199, 16203
; OTHER INFORMATION: n = g or t
; NAME/KEY: misc.feature
; LOCATION: 3619
; OTHER INFORMATION: n = a or t
; NAME/KEY: misc.feature
; LOCATION: 14547
; OTHER INFORMATION: nucleotide in position 14547 is t, or absent
US-09-962-677-8

```

```

Query Match      68.3%; Score 811; DB 36; Length 18597;
Best Local Similarity 99.5%; Pred. No. 8,8e-93;
Matches 1181; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

Qy 1 GATCGCCCACTGCACTCAGCCTGGTGGAGAGACGAGACTTGTTCAAAAAAAAAA 60
Db 78 GATCGCCCACTGCACTCAGCCTGGTGGAGAGACGAGACTTGTTCAAAAAAAAAA 137

```


198 TTTTCTTTTCTTTTCTTTTGGGACAGTCTGCTCTGCTGCGCCAGGCTGAGTACAATGGT 257
Db
181 CGGATCTTGGCTCACTGCAACCTCTGCTCCAGGTTCAAGAAATCTTCTGCTCAGCC 240
Qy
258 CGGATCTTGGCTCACTGCAACCTCTGCTCCAGGTTCAAGAAATCTTCTGCTCAGCC 317
Db
241 TCCCAAGTACCAACACGCGGCTGAGTATTTTGTATTTTGTAGAGAGGGGTTTCA 300
Qy
318 TCCCAAGTACCAACACGCGGCTGAGTATTTTGTATTTTGTAGAGAGGGGTTTCA 377
Db
301 CATGTTGTCCAGGCTGCTGCTGAGTCTCTGACCTCAGGTGATCCACCGGCTCGGCCCC 360
Qy
378 CATGTTGTCCAGGCTGCTGCTGAGTCTCTGACCTCAGGTGATCCACCGGCTCGGCCCC 437
Db
361 CAAAGTACTAGGATTAACAGGCTGAGCAGCGCTGCGGCTGCGGCTTTTAAATCA 420
Qy
438 CAAAGTACTAGGATTAACAGGCTGAGCAGCGCTGCGGCTGCGGCTTTTAAATCA 497
Db
421 AGTAAAGAGCTCAATTAACAGGCTGAGCAGCGCTGCGGCTGCGGCTTTTAAATCA 480
Qy
498 AGTAAAGAGCTCAATTAACAGGCTGAGCAGCGCTGCGGCTGCGGCTTTTAAATCA 556
Db
481 AAATGCAAAATCCCTTATTTAGTTTGTAGGAAACAGATCTCAAAACAGCAGTTTGTGACAAG 540
Qy
557 AAATGCAAAATCCCTTATTTAGTTTGTAGGAAACAGATCTCAAAACAGCAGTTTGTGACAAG 616
Db
541 ACCGACGAGAAACGTTGGGCTGAGCAGCGCTGCGGCTGCGGCTTTTAAATCA 600
Qy
617 ACCGACGAGAAACGTTGGGCTGAGCAGCGCTGCGGCTGCGGCTTTTAAATCA 676
Db
601 TTCCCAAGAGGGGCTGAGTATTTTGTATTTTGTAGAGAGGGGTTTCA 660
Qy
677 TTCCCAAGAGGGGCTGAGTATTTTGTATTTTGTAGAGAGGGGTTTCA 736
Db
661 AAGACTCTCAGCTGTGGGCTGAGTATTTTGTATTTTGTAGAGAGGGGTTTCA 720
Qy
737 AAGACTCTCAGCTGTGGGCTGAGTATTTTGTATTTTGTAGAGAGGGGTTTCA 796
Db
721 CCCCTGGGCGACGCTCTCTAGAGCGGGGCTGCGGCTGCGGCTTTTAAATCA 780
Qy

Query Match 55.4%; Score 658; DB 22; Length 18596;
Best Local Similarity 99.2%; Pred. No. 7.7e-74;
Matches 1178; Conservative 0; Mismatches 8; Indels 1; Gaps 1;
Qy 1 GATCGCGCACTGCACTCCAGGCTGGGTGAGAGCGAGGCTCTGCTCAAAAAA 60
Db 78 GATCGCGCACTGCACTCCAGGCTGGGTGAGAGCGAGGCTCTGCTCAAAAAA 137
Qy 61 AAAAAAGCGCGGCTGCACTCCAGGCTGGGTGAGAGCGAGGCTCTGCTCAAAAAA 120
Db 138 AAAAAAGCGCGGCTGCACTCCAGGCTGGGTGAGAGCGAGGCTCTGCTCAAAAAA 197
Qy 121 TTTTCTTTTCTTTTCTTTTGGGACAGTCTGCTGCTGCGCCAGGCTGGAGTACAATGGT 180
Db 198 TTTTCTTTTCTTTTCTTTTGGGACAGTCTGCTGCTGCGCCAGGCTGGAGTACAATGGT 257
Qy 181 CGGATCTTGGCTCACTGCAACCTCTGCTCCAGGTTCAAGCAATCTTCTGCTCAGCC 240
Db 258 CGGATCTTGGCTCACTGCAACCTCTGCTCCAGGTTCAAGCAATCTTCTGCTCAGCC 317
Qy 241 TCCCAAGTACCAACACGCGGCTGAGTATTTTGTATTTTGTAGAGAGGGGTTTCA 300
Db 318 TCCCAAGTACCAACACGCGGCTGAGTATTTTGTATTTTGTAGAGAGGGGTTTCA 377
Qy 301 CATGTTGTCCAGGCTGCTGCTGAGTCTCTGACCTCAGGTGATCCACCGGCTCGGCCCC 360
Db 378 CATGTTGTCCAGGCTGCTGCTGAGTCTCTGACCTCAGGTGATCCACCGGCTCGGCCCC 437
Qy 361 CAAAGTACTAGGATTAACAGGCTGAGCAGCGCTGCGGCTGCGGCTTTTAAATCA 420
Db 438 CAAAGTACTAGGATTAACAGGCTGAGCAGCGCTGCGGCTGCGGCTTTTAAATCA 497
Qy 421 AGTAAAGAGCTCAATTAACAGGCTGAGCAGCGCTGCGGCTGCGGCTTTTAAATCA 480
Db 498 AGTAAAGAGCTCAATTAACAGGCTGAGCAGCGCTGCGGCTGCGGCTTTTAAATCA 556
Qy 481 AAATGCAAAATCCCTTATTTAGTTTGTAGGAAACAGATCTCAAAACAGCAGTTTGTGACAAG 540
Db 557 AAATGCAAAATCCCTTATTTAGTTTGTAGGAAACAGATCTCAAAACAGCAGTTTGTGACAAG 616
Qy 541 ACCGACGAGAAACGTTGGGCTGAGCAGCGCTGCGGCTGCGGCTTTTAAATCA 600
Db 617 ACCGACGAGAAACGTTGGGCTGAGCAGCGCTGCGGCTGCGGCTTTTAAATCA 676
Qy 601 TTCCCAAGAGGGGCTGAGTATTTTGTATTTTGTAGAGAGGGGTTTCA 660
Db 677 TTCCCAAGAGGGGCTGAGTATTTTGTATTTTGTAGAGAGGGGTTTCA 736
Qy 661 AAGACTCTCAGCTGTGGGCTGAGTATTTTGTATTTTGTAGAGAGGGGTTTCA 720
Db 737 AAGACTCTCAGCTGTGGGCTGAGTATTTTGTATTTTGTAGAGAGGGGTTTCA 796
Qy 721 CCCCTGGGCGACGCTCTCTAGAGCGGGGCTGCGGCTGCGGCTTTTAAATCA 780
Db

198 TTTTCTTTTCTTTTCTTTTGGGACAGTCTGCTCTGCTGCGCCAGGCTGAGTACAATGGT 257
Db
181 CGGATCTTGGCTCACTGCAACCTCTGCTCCAGGTTCAAGAAATCTTCTGCTCAGCC 240
Qy
258 CGGATCTTGGCTCACTGCAACCTCTGCTCCAGGTTCAAGAAATCTTCTGCTCAGCC 317
Db
241 TCCCAAGTACCAACACGCGGCTGAGTATTTTGTATTTTGTAGAGAGGGGTTTCA 300
Qy
318 TCCCAAGTACCAACACGCGGCTGAGTATTTTGTATTTTGTAGAGAGGGGTTTCA 377
Db
301 CATGTTGTCCAGGCTGCTGCTGAGTCTCTGACCTCAGGTGATCCACCGGCTCGGCCCC 360
Qy
378 CATGTTGTCCAGGCTGCTGCTGAGTCTCTGACCTCAGGTGATCCACCGGCTCGGCCCC 437
Db
361 CAAAGTACTAGGATTAACAGGCTGAGCAGCGCTGCGGCTGCGGCTTTTAAATCA 420
Qy
438 CAAAGTACTAGGATTAACAGGCTGAGCAGCGCTGCGGCTGCGGCTTTTAAATCA 497
Db
421 AGTAAAGAGCTCAATTAACAGGCTGAGCAGCGCTGCGGCTGCGGCTTTTAAATCA 480
Qy
498 AGTAAAGAGCTCAATTAACAGGCTGAGCAGCGCTGCGGCTGCGGCTTTTAAATCA 556
Db
481 AAATGCAAAATCCCTTATTTAGTTTGTAGGAAACAGATCTCAAAACAGCAGTTTGTGACAAG 540
Qy
557 AAATGCAAAATCCCTTATTTAGTTTGTAGGAAACAGATCTCAAAACAGCAGTTTGTGACAAG 616
Db
541 ACCGACGAGAAACGTTGGGCTGAGCAGCGCTGCGGCTGCGGCTTTTAAATCA 600
Qy
617 ACCGACGAGAAACGTTGGGCTGAGCAGCGCTGCGGCTGCGGCTTTTAAATCA 676
Db
601 TTCCCAAGAGGGGCTGAGTATTTTGTATTTTGTAGAGAGGGGTTTCA 660
Qy
677 TTCCCAAGAGGGGCTGAGTATTTTGTATTTTGTAGAGAGGGGTTTCA 736
Db
661 AAGACTCTCAGCTGTGGGCTGAGTATTTTGTATTTTGTAGAGAGGGGTTTCA 720
Qy
737 AAGACTCTCAGCTGTGGGCTGAGTATTTTGTATTTTGTAGAGAGGGGTTTCA 796
Db
721 CCCCTGGGCGACGCTCTCTAGAGCGGGGCTGCGGCTGCGGCTTTTAAATCA 780
Qy

RESULT 10
US-09-631-275-140
; Sequence 140, Application US/09631275
; GENERAL INFORMATION:
; APPLICANT: Hong Chen
; APPLICANT: Nelson B. Freimer
; APPLICANT: Thomas Novak
; APPLICANT: Thomas Novak

Chromosome-18
for Disclosing and Treating Chromosome-18

RESULT 11
US-09-880-107-1590
Sequence 1590, Application US/09880107
GENERAL INFORMATION:
APPLICANT: Horne, Darci T.
APPLICANT: Vockley, Joseph G.
APPLICANT: Scherf, Uwe
APPLICANT: Gene Logic, Inc.
TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
FILE REFERENCE: 44921-5028-WO
CURRENT APPLICATION NUMBER: US/09/880,107
CURRENT FILING DATE: 2001-06-14
PRIOR APPLICATION NUMBER: US 60/211,379
PRIOR FILING DATE: 2000-06-14
PRIOR APPLICATION NUMBER: US 60/237,054
PRIOR FILING DATE: 2000-10-02
NUMBER OF SEQ ID NOS: 3950
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1590
LENGTH: 18596
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Genbank Accession No. D00596
US-09-880-107-1590
Query Match 55.4%; Score 658; DB 33; Length 18596;
Best Local Similarity 99.2%; Pred. No. 7.7e-74;
Matches 1178; Conservative 0; Mismatches 8; Indels 1; Gaps 1;
QY 1 GATCGCGCCACTGCTCCAGCTCCAGCTCGGTGAGAGCGAGACTCTGTCTCAAAAAA 60
DB 78 GATCGCGCCACTGCTCCAGCTCCAGCTCGGTGAGAGCGAGACTCTGTCTCAAAAAA 137
QY 61 AAAAAAGCGCGAGGCTCAACAAAAAACCTCGGAAAGCCCTGCGGTCTTTTTTTT 120
DB 138 AAAAAAGCGCGAGGCTCAACAAAAAACCTCGGAAAGCCCTGCGGTCTTTTTTTT 197
QY 121 TTTTTTTTTTTTTTTTTTGGGACGCTTGTCTGTCGCCAGGCTGAGTACAAATGGT 180
DB 198 TTTTTTTTTTTTTTTTTTGGGACGCTTGTCTGTCGCCAGGCTGAGTACAAATGGT 257
QY 181 CGGACTTGCTCACTCAACCTTCCCTCCAGCTTCAAGCAATTTCTGCTCAGCC 240
DB 258 CGGACTTGCTCACTCAACCTTCCCTCCAGCTTCAAGCAATTTCTGCTCAGCC 317
QY 241 TCCCAAGTAGCCACACGCCAGCTAATTTTTGTATTTTGTAGTACAGCGGGTTTAC 300
DB 318 TCCCAAGTAGCCACACGCCAGCTAATTTTTGTATTTTGTAGTACAGCGGGTTTAC 377
QY 301 CATGTTGCCAGGCTGTTCTNGAATCTTGAATCACTGAGTATCAACCGCTCGGCCCC 360
DB 378 CATGTTGCCAGGCTGTTCTNGAATCTTGAATCACTGAGTATCAACCGCTCGGCCCC 437
QY 361 CAAAGTACTAGGATTAAGGCTGAGCCACCGGTCAGCGCCCTGCGGTTTAAATCA 420
DB 438 CAAAGTACTAGGATTAAGGCTGAGCCACCGGTCAGCGCCCTGCGGTTTAAATCA 497
QY 421 AGTAGAAAAGCTGATTAATACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
DB 498 AGTAGAAAAGCTGATTAATACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 556
QY 481 AATGCAATCTTATTTAGTTTGTAGGAAACAGATCTCAACAGCAGTTTGTGNGACAAG 540
DB 557 AATGCAATCTTATTTAGTTTGTAGGAAACAGATCTCAACAGCAGTTTGTGNGACAAG 616
QY 541 ACCGAGGAAACCTGGAATCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
DB 617 ACCGAGGAAACCTGGAATCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 676
QY 601 TTTCCAAAGCGGCTGCTTCCGCGCCACCGCACTGCTGCTGCTGCTGCTGCTGCTGCT 660
DB 677 TTTCCAAAGCGGCTGCTTCCGCGCCACCGCACTGCTGCTGCTGCTGCTGCTGCTGCT 736

QY 661 AAGACTCTCAGCTGTGGCCCTGGCTCCGTTCTGTGTCACACCCGCTGCTCTTCTTC 720
DB 737 AAGACTCTCAGCTGTGGCCCTGGCTCCGTTCTGTGTCACACCCGCTGCTCTTCTTC 796
QY 721 CCCCTGGCCACGCTCTCTAGAGCGGGGCGCGCCGAGACCCCGCGAGAGAGGCG 780
DB 797 CCCCTGGCCACGCTCTCTAGAGCGGGGCGCGCCGAGACCCCGCGAGAGAGGCG 856
QY 781 GAGCGGGGAGCGCGCGGAGAAAGGCGCGCGAGAGGGGTCTGTCACCGCGCCACTTGG 840
DB 857 GAGCGGGGAGCGCGCGGAGAAAGGCGCGCGAGAGGGGTCTGTCACCGCGCCACTTGG 916
QY 841 CCTGCTTCCGTCGCGCGCGCACTTGGCTGCTGCTCCGTCGCTCCGTCGCGCGCTG 900
DB 917 CCTGCTTCCGTCGCGCGCGCACTTGGCTGCTGCTCCGTCGCTCCGTCGCGCGCTG 976
QY 901 CCTCGCTCCGTCGCGCGCGCACTTGGCTGCTGCTCCGTCGCGCGCTGCGAGCTGCGCGCGCC 960
DB 977 CCTCGCTCCGTCGCGCGCGCACTTGGCTGCTGCTCCGTCGCTCCGTCGCGCGCGCC 1036
QY 961 TTGCCCCCGCCGACAGAGCGGAGCGCGAGCGCGCTCCGCGCACCGGAGCTGCGAG 1020
DB 1037 TTGCCCCCGCCGACAGAGCGGAGCGCGAGCGCGCTCCGCGCACCGGAGCTGCGAG 1096
QY 1021 TACTTGGGAGATCCAAACACATCTCTCGCTGCGCGCTCAGGAAAGACGACCGCACCGGC 1080
DB 1097 TACTTGGGAGATCCAAACACATCTCTCGCTGCGCGCTCAGGAAAGACGACCGCACCGGC 1156
QY 1081 ACCGCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
DB 1157 ACCGCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1216
QY 1141 GCGCCCTCGCGGAGCGGTGCGGAGAGGAGGAGCGCGCTGGGGA 1187
DB 1217 GCGCCCTCGCGGAGCGGTGCGGAGAGGAGGAGCGCGCTGGGGA 1263

RESULT 12
US-09-954-531-124
Sequence 124, Application US/09954531
GENERAL INFORMATION:
APPLICANT: Weaver, Zoe
TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc
TITLE OF INVENTION: Gene Sets
FILE REFERENCE: 689290-77
CURRENT APPLICATION NUMBER: US/09/954,531
CURRENT FILING DATE: 2002-05-02
PRIOR APPLICATION NUMBER: US/60/233,133
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: US/60/234,009
PRIOR FILING DATE: 2000-09-20
PRIOR APPLICATION NUMBER: US/60/234,034
PRIOR FILING DATE: 2000-09-20
PRIOR APPLICATION NUMBER: US/60/234,509
PRIOR FILING DATE: 2000-09-22
PRIOR APPLICATION NUMBER: US/60/234,567
PRIOR FILING DATE: 2000-09-22
NUMBER OF SEQ ID NOS: 1392
SOFTWARE: PatentIn version 3.0
SEQ ID NO 124
LENGTH: 18596
TYPE: DNA
ORGANISM: Homo sapiens
US-09-954-531-124
Query Match 55.4%; Score 658; DB 36; Length 18596;
Best Local Similarity 99.2%; Pred. No. 7.7e-74;
Matches 1178; Conservative 0; Mismatches 8; Indels 1; Gaps 1;
QY 1 GATCGCGCCACTGCTCCAGCTCCAGCTCGGTGAGAGCGAGACTCTGTCTCAAAAAA 60
DB 78 GATCGCGCCACTGCTCCAGCTCCAGCTCGGTGAGAGCGAGACTCTGTCTCAAAAAA 137

QY	1	GATCGCCCACTGCACTCAGCCTGGGTGAGAGCGAGCTCTCTCTCAAAAAA	60
Db	78	GATCGCCCACTGCACTCAGCCTGGGTGAGAGCGAGCTCTCTCTCAAAAAA	137
QY	61	AAAAAGACCCGACGAGCTCAAAACAAAAA	120
Db	138	AAAAAGACCCGACGAGCTCAAAACAAAAA	197
QY	121	TT	180
Db	198	TT	257
QY	181	CGGATCTTGGCTACTGCAACCTCCAGCTTCAAGCAATCTTCTGCTCAGCC	240
Db	258	CGGATCTTGGCTACTGCAACCTCCAGCTTCAAGCAATCTTCTGCTCAGCC	317
QY	241	TCCCAAGTAGCCACCGCCAGCTAAATTTTGTANTTTTAGTAGAGCGGGTTTCA	300
Db	318	TCCCAAGTAGCCACCGCCAGCTAAATTTTGTANTTTTAGTAGAGCGGGTTTCA	377
QY	301	CATGTTGTCAGGCTGCTCTGAACTCTCTGACCTCAGTATCAACCGCTCGCC	360
Db	378	CATGTTGTCAGGCTGCTCTGAACTCTCTGACCTCAGTATCAACCGCTCGCC	437
QY	361	CAAGTACTTAGGATTAACAGCGTACGCAACCGCTCAGCGCTTGTAAATCA	420
Db	438	CAAGTACTTAGGATTAACAGCGTACGCAACCGCTCAGCGCTTGTAAATCA	497
QY	421	AGTAGAAAAGCTGATTAACCACTTCTGCTGCTGCTGCTGCTGCTGCTGCT	480
Db	498	AGTAGAAAAGCTGATTAACCACTTCTGCTGCTGCTGCTGCTGCTGCTGCT	556
QY	481	AAATGCAATCCTTTTATTTAGTTAGTAAACAGATCTCAAAACAGCAGTTT	540
Db	557	AAATGCAATCCTTTTATTTAGTTAGTAAACAGATCTCAAAACAGCAGTTT	616
QY	541	ACGCGAGAAAACGTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	600
Db	617	ACGCGAGAAAACGTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	676
QY	601	TTCCCAAGGCGCAGTCTCTCCCGCCACCGCACTGCTGCTGCTGCTGCTGCT	660
Db	677	TTCCCAAGGCGCAGTCTCTCCCGCCACCGCACTGCTGCTGCTGCTGCTGCT	736
QY	661	AAGACTCTCAGCTGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	720
Db	737	AAGACTCTCAGCTGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	796
QY	721	CCCTGCGCAGCCTCTTAGAGCGGGGCGCGGACCCCGCGAGCAGAGAGCGG	780
Db	797	CCCTGCGCAGCCTCTCTTAGAGCGGGGCGCGGACCCCGCGAGCAGAGAGCGG	856
QY	781	GAGCGGCGGAGCGCGCGGAGAAAGCGCGGAGAGGGGCTCTGCGCAGCGCC	840
Db	857	GAGCGGCGGAGCGCGCGGAGAAAGCGCGGAGAGGGGCTCTGCGCAGCGCC	916
QY	841	CCCTGCTCCGCTGCGCGCGCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTG	900
Db	917	CCCTGCTCCGCTGCGCGCGCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTG	976
QY	901	CTTCCGTCCTCCCGCGCGCGCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTG	960
Db	977	CTTCCGTCCTCCCGCGCGCGCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTG	1036
QY	961	TTGCCCCCGCGCGCGCGCGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	1020
Db	1037	TTGCCCCCGCGCGCGCGCGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	1096
QY	1021	TACCTGGGCGCAGATCCAAACATCTCTCGCTGCTGCTGCTGCTGCTGCTG	1080
Db	1097	TACCTGGGCGCAGATCCAAACATCTCTCGCTGCTGCTGCTGCTGCTGCTG	1156
QY	1081	ACCGGCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1140

QY	541	ACCAGAGAAAACGTTGGGAACTGTGCTGCTGCTGCTTAGAGAGCGCGCTGACAGCGG	600
Db	617	ACCAGAGAAAACGTTGGGAACTGTGCTGCTGCTTAGAGAGCGCGCTGACAGAGCGG	676
QY	601	TTCCCAAGGCGCAGTCTCTCCCGCCACCGCACTGCTGCTGCTGCTGCTGCTGCT	660
Db	677	TTCCCAAGGCGCAGTCTCTCCCGCCACCGCACTGCTGCTGCTGCTGCTGCTGCT	736
QY	661	AAGACTCTCAGCTGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	720
Db	737	AAGACTCTCAGCTGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	796
QY	721	CCCTGCGCAGCCTCTTAGAGCGGGGCGCGGACCCCGCGAGCAGAGAGAGCGG	780
Db	797	CCCTGCGCAGCCTCTCTTAGAGCGGGGCGCGGACCCCGCGAGCAGAGAGAGCGG	856
QY	781	GAGCGGCGGAGCGCGCGGAGAAAGCGCGGAGAGGGTCTTGCACCGCGCAGTTGG	840
Db	857	GAGCGGCGGAGCGCGCGGAGAAAGCGCGGAGAGGGTCTTGCACCGCGCAGTTGG	916
QY	841	CCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	900
Db	917	CCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	976
QY	901	CCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	960
Db	977	CCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1036
QY	961	TTGCCCCCGCGCGCGCGCGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	1020
Db	1037	TTGCCCCCGCGCGCGCGCGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	1096
QY	1021	TACCTGGGCGCAGATCCAAACATCTCTCGCTGCTGCTGCTGCTGCTGCTGCTG	1080
Db	1097	TACCTGGGCGCAGATCCAAACATCTCTCGCTGCTGCTGCTGCTGCTGCTGCTG	1156
QY	1081	ACCGGCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1140
Db	1157	ACCGGCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1216
QY	1141	GGCCCCCTGCGGAGCGGTTGGGAGAGAGGAGGCGCGGCTGGGGA	1187
Db	1217	GGCCCCCTGCGGAGCGGTTGGGAGAGAGGAGGCGCGGCTGGGGA	1263

RESULT 14
US-09-967-768A-119
Sequence 119, Application US/09967768A
GENERAL INFORMATION:
APPLICANT: Augustus, Meena
TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
FILE REFERENCE: 689290-72
CURRENT APPLICATION NUMBER: US/09/967,768A
CURRENT FILING DATE: 2001-09-28
PRIOR APPLICATION NUMBER: US/60/236,109
PRIOR FILING DATE: 2000-09-28
PRIOR APPLICATION NUMBER: US/60/236,034
PRIOR FILING DATE: 2000-09-28
PRIOR APPLICATION NUMBER: US/60/236,111
PRIOR FILING DATE: 2000-09-28
NUMBER OF SEQ ID NOS: 325
SOFTWARE: PatentIn version 3.0
SEQ ID NO 119
LENGTH: 18596
TYPE: DNA
ORGANISM: Homo sapiens
US-09-967-768A-119
Query Match 55.4%; Score 658; DB 36; Length 18596;
Best Local Similarity 99.2%; Pred. No. 7.7e-74;
Matches 1178; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

RESULT 15

25.0%: Score 297; DB 1; Length 1536;

QY 833 CCAATTGGCGTGGCTCCGTCCTCCGCGCGGCCACTTGGCGTGGCTCCGTCCTCCGCGCGGCCAC 892

D6 14 CCACTTGGCGTGGCTCCGTCCTCCGCGCGGCCACTTGGCGTGGCTCCGTCCTCCGCGCGGCCAC 73

db 14 CCACTTGGCCCTCCGTCCCGCGGCCACTTGGCCTGCCTCCGTCCCGCGGCCAC 73

952

QY
893 TTGGCCTGCCCTCGGTCCCGCCGCGCCATGCTTGTCGCGGCACGAGC

74 TTATGCGCTGCGTCCCCCGCCGCGCATGCCCTGTGGCCGGCTCGGAGCTGCCGC 133

BB
J#
1012

QY
953 GCCGGCCCTTGCCCCCCCCCGCCGACAGGAGCGGGACGCCGAGCCGGGTCCGCCGACGGGG
101

193

[illegible]

1013 AGCTGCAGTACCTGGGGCAGATCCAACACATCCTCGCTGCGCGTCAGGAAGGACGACC 1072

253

Db 194 AGCTGCAGTACCTGGGGCAGATCCAAACACATCCCGCTGCGGGCTGACGCTGCGGCTG

1073 GCACGGGCACCGGCACCCCTGTCGGTATTCGGCATGCAGGCGGCTACAGCCTGAGAG 1129

CCCTGAGAG 310

db 254 GCACGGGCACGGCACCCCTGTCCGTAATTCGGCATGCGAGGGCGCGCTACAGCCCTGAGAGC

Search completed: November 29, 2002, 05:05:02
Job time : 2485.99 secs

[illegible]

Y 1082 CCGCACCTCTCGGTATCGGATCGAGCGCGCTACAGCTGAG 1129
b 61 CCGCACCTCTCGGTATCGGATCGAGCGCGCTACAGCTGAG 108

RESULT 6

US-09-764-869-2113/c
; Sequence 2113, Application US/09764869
; Patent No. US20020061521A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007
; CURRENT APPLICATION NUMBER: US/09/764,869
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2113
; LENGTH: 10867
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-869-2113

Query Match 4.9%; Score 58; DB 10; Length 10867;
Best Local Similarity 100.0%; Pred. No. 2.8e-08;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 GCACCTGCACTCCAGCTCGGTGAGAGCGAGACTCTGTCTCAAAAAA 64
Db 5892 GCCACTGCACTCCAGCTCGGTGAGAGCGAGACTCTGTCTCAAAAAA 5835

RESULT 7

US-09-764-877-3970/c
; Sequence 3970, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3970
; LENGTH: 10867
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-3970

Query Match 4.9%; Score 58; DB 10; Length 10867;
Best Local Similarity 100.0%; Pred. No. 2.8e-08;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 GCACCTGCACTCCAGCTCGGTGAGAGCGAGACTCTGTCTCAAAAAA 64
Db 5892 GCCACTGCACTCCAGCTCGGTGAGAGCGAGACTCTGTCTCAAAAAA 5835

RESULT 8

US-09-764-869-2112/c
; Sequence 2112, Application US/09764869
; Patent No. US20020061521A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007
; CURRENT APPLICATION NUMBER: US/09/764,869
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper

; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2112
; LENGTH: 10907
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-869-2112

Query Match 4.9%; Score 58; DB 10; Length 10907;
Best Local Similarity 100.0%; Pred. No. 2.8e-08;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 GCACCTGCACTCCAGCTCGGTGAGAGCGAGACTCTGTCTCAAAAAA 64
Db 5929 GCCACTGCACTCCAGCTCGGTGAGAGCGAGACTCTGTCTCAAAAAA 5872

RESULT 9

US-09-764-877-3968/c
; Sequence 3968, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3968
; LENGTH: 10907
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-3968

Query Match 4.9%; Score 58; DB 10; Length 10907;
Best Local Similarity 100.0%; Pred. No. 2.8e-08;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 GCACCTGCACTCCAGCTCGGTGAGAGCGAGACTCTGTCTCAAAAAA 64
Db 5929 GCCACTGCACTCCAGCTCGGTGAGAGCGAGACTCTGTCTCAAAAAA 5872

RESULT 10

US-09-764-877-3972/c
; Sequence 3972, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3972
; LENGTH: 12718
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-3972

Query Match 4.9%; Score 58; DB 10; Length 12718;
Best Local Similarity 100.0%; Pred. No. 2.6e-08;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 GCACCTGCACTCCAGCTCGGTGAGAGCGAGACTCTGTCTCAAAAAA 64
Db 5929 GCCACTGCACTCCAGCTCGGTGAGAGCGAGACTCTGTCTCAAAAAA 5872


```
RESULT 11
US-09-729-920-3/c
; Sequence 3, Application US/09729920
; Patent No. US20020103115A1
; GENERAL INFORMATION:
; APPLICANT: GUEGLER, Karl et al.
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL000858
; CURRENT APPLICATION NUMBER: US/09/729,920
; CURRENT FILING DATE: 2000-12-06
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 143306
; TYPE: DNA
; ORGANISM: Human
US-09-729-920-3

Query Match      4.7%; Score 56; DB 10; Length 143306;
Best Local Similarity 100.0%; Pred. No. 4.4e-08;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 189 GGCCTACTGCAACCTCTGCTCCAGGTTCAAGCAATTTCTTGCTCAGCTCC 244
      |||||||
Db 65791 GGCCTACTGCAACCTCTGCTCCAGGTTCAAGCAATTTCTTGCTCAGCTCC 65736

RESULT 12
US-09-764-869-2111/c
; Sequence 2111, Application US/09764869
; Patent No. US20020061521A1
; GENERAL INFORMATION:
; APPLICANT: ROSEN ET AL.
; TITLE OF INVENTION: NUCLEIC ACIDS, PROTEINS, AND ANTIBODIES
; FILE REFERENCE: PC007
; CURRENT APPLICATION NUMBER: US/09/764,869
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2111
; LENGTH: 10894
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-869-2111

Query Match      4.6%; Score 55; DB 10; Length 10894;
Best Local Similarity 100.0%; Pred. No. 1.7e-07;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 GCCACTGCACTCCAGCTGGGTGAGAGCGAGACTCTGTCTCAAAAAA 61
      |||||||
Db 5911 GCCACTGCACTCCAGCTGGGTGAGAGCGAGACTCTGTCTCAAAAAA 5857

RESULT 13
US-09-982-091A-5
; Sequence 5, Application US/09982091A
; Patent No. US20020151030A1
; GENERAL INFORMATION:
; APPLICANT: CALIFORNIA INSTITUTE OF TECHNOLOGY
; APPLICANT: KUMAGAI, AKIKO
; APPLICANT: DUNPHY, WILLIAM
; TITLE OF INVENTION: CLASPIN PROTEINS AND METHODS OF USE THEREOF
; FILE REFERENCE: CIT1320-1
; CURRENT APPLICATION NUMBER: US/09/982,091A
; CURRENT FILING DATE: 2002-10-17
; PRIOR APPLICATION NUMBER: US 60/241,246
; PRIOR FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
```

```
; SEQ ID NO 5
; LENGTH: 58837
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-982-091A-5

Query Match      4.5%; Score 54; DB 10; Length 58837;
Best Local Similarity 100.0%; Pred. No. 1.9e-07;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 TGCATCCAGCTGGGTGAGAGCGAGACTCTGTCTCAAAAAA 65
      |||||||
Db 14914 TGCATCCAGCTGGGTGAGAGCGAGACTCTGTCTCAAAAAA 14967

RESULT 14
US-09-764-860-1108
; Sequence 1108, Application US/09764860
; Patent No. US20020094953A1
; GENERAL INFORMATION:
; APPLICANT: ROSEN ET AL.
; TITLE OF INVENTION: NUCLEIC ACIDS, PROTEINS, AND ANTIBODIES
; FILE REFERENCE: PC008
; CURRENT APPLICATION NUMBER: US/09/764,860
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1198
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1108
; LENGTH: 32170
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-860-1108

Query Match      4.0%; Score 48; DB 10; Length 32170;
Best Local Similarity 100.0%; Pred. No. 9.2e-06;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 197 GCAACCTCTGCTCCAGGTTCAAGCAATTTCTTGCTCAGCTCCC 244
      |||||||
Db 3592 GCAACCTCTGCTCCAGGTTCAAGCAATTTCTTGCTCAGCTCCC 3639

RESULT 15
US-09-954-531-145/c
; Sequence 145, Application US/09954531
; Patent No. US20020165180A1
; GENERAL INFORMATION:
; APPLICANT: WEAVER, ZOE
; TITLE OF INVENTION: PROCESS FOR IDENTIFYING ANTI-CANCER THERAPEUTIC AGENTS USING C
; FILE REFERENCE: 689290-77
; CURRENT APPLICATION NUMBER: US/09/954,531
; CURRENT FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: US/60/233,133
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,009
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,034
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,509
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US/60/234,567
; NUMBER OF SEQ ID NOS: 1392
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 145
; LENGTH: 13500
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(13500)
```

; OTHER INFORMATION: n=a,t,g or c
JS-09-954-531-145

Query Match 3.9%; Score 46; DB 9; Length 13500;
Best Local Similarity 100.0%; Pred. No. 4e-05;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2Y 184 ATCTTGGCTCACTGCAACCTCTGCTCCCGGTTCAAGCAATTCTT 229
|||||
Db 9584 ATCTTGGCTCACTGCAACCTCTGCTCCCGGTTCAAGCAATTCTT 9539

Search completed: November 29, 2002, 05:51:16
Job time : 584.685 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

DN nucleic - nucleic search, using sw model

Run on: November 29, 2002, 00:20:35 ; Search time 48.3867 Seconds
(without alignments)
7523.251 Million cell updates/sec

Title: US-09-963-333-7
Perfect score: 1187
Sequence: 1 gatcgccgactgcactcca.....gagggagcgcgctgggga 1187

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 441362 seqs, 153338381 residues

Word size : 15

Total number of hits satisfying chosen parameters: 23436

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/2/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	658	55.4	18596	4	US-09-318-448-11
2	297	25.0	1536	3	US-09-089-195-1
3	297	25.0	1536	4	US-09-367-007C-38
4	205	17.3	942	4	US-09-347-878-29
5	46	3.9	72604	4	US-09-268-992-7
6	46	3.9	72604	4	US-09-657-474-7
7	40	3.4	2900	3	US-09-038-832-1
8	40	3.4	4136	4	US-09-103-875-2
9	40	3.4	35100	1	PCT-US93-06251-19
10	40	3.4	35100	5	US-08-306-691B-19
11	39	3.3	1014	4	US-09-257-179-32
12	39	3.3	1386	2	US-08-667-080-76
13	39	3.3	5300	4	US-08-938-669A-1
14	39	3.3	5304	4	US-08-938-669A-2
15	39	3.3	6169	4	US-08-938-669A-3
16	39	3.3	72604	4	US-09-268-992-7
17	39	3.3	72604	4	US-09-657-474-7
18	39	3.3	112132	4	US-09-741-150-3
19	39	3.3	176373	3	US-09-128-155-17
20	38	3.2	38844	3	US-09-734-675-3
21	38	3.2	70000	4	US-09-851-896-3
22	38	3.2	162450	4	US-09-345-888-1
23	37	3.1	6792	4	US-09-374-454-20
24	37	3.1	43950	4	US-09-735-934A-3
25	36	3.0	461	4	US-09-404-879A-1
26	36	3.0	461	4	US-09-404-879A-3
27	36	3.0	589	4	US-09-385-982-455

ALIGNMENTS

RESULT 1

US-09-318-448-11

; Patent No. 6210350

; APPLICANT: Johnson, William G.

; APPLICANT: Stearros, Edward S.

; TITLE OF INVENTION: METHODS FOR DIAGNOSING, PREVENTING, AND TREATING

; TITLE OF INVENTION: DEVELOPMENTAL DISORDERS

; FILE REFERENCE: 601-1-057

; CURRENT APPLICATION NUMBER: US/09/318,448

; CURRENT FILING DATE: 1999-05-25

; NUMBER OF SEQ ID NOS: 46

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 11

; LENGTH: 18596

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-318-448-11

Query Match	55.4%	Score 658;	DB 4;	Length 18596;
Best Local Similarity	99.2%;	Pred. No. 4.1e-215;		
Matches 1178;	Conservative	0;	Mismatches 8;	Indels 1;
Gaps	1;			
QY	1	GATCGCGCCACTGCACCTCCAGCCTGGGTGAGAGAGCGAGACTCTGTCTCAAAAAAAAAA	60	
DB	78	GATCGCGCCACTGCACCTCCAGCCTGGGTGAGAGAGCGAGACTCTGTCTCAAAAAAAAAA	137	
QY	61	AAAGAGCGCCAGGCTCAACAAAAAACCCTCGAAAGCCCTGGGGCTCTTTTTTTT	120	
DB	138	AAAAAGCGCGAGGCTCAACAAAAAACCCTCGAAAGCCCTGGGGCTCTTTTTTTT	197	
QY	121	TTTTTTTTTTTTTTTTTTTGGGACAGTCTGTCTGTGCGCCAGGCTGAGTACATCGT	180	
DB	198	TTTTTTTTTTTTTTTTTTTGGGACAGTCTGTCTGTGCGCCAGGCTGAGTACATCGT	257	
QY	181	CGGATCTTGGCTCACTGCAACCTCTGCTCCAGGTTCAAGCAATTCCTTCCTCCAGCC	240	
DB	258	CGGATCTTGGCTCACTGCAACCTCTGCTCCAGGTTCAAGCAATTCCTTCCTCCAGCC	317	
QY	241	TCCCAAGTAGCACACGCCAGCTAAATTTGTGTTAGTAGAGCGGGGTTTCAC	300	
DB	318	TCCCAAGTAGCACACGCCAGCTAAATTTGTGTTAGTAGAGCGGGGTTTCAC	377	
QY	301	CATGTTGTCAGGCTGCTGACCTCTGACCTCAGTGATCCACCCGCTCGGCCCCC	360	
DB	378	CATGTTGTCAGGCTGCTGACCTCTGACCTCAGTGATCCACCCGCTCGGCCCCC	437	
QY	361	CBAAGTACTAGGATTACAGCGGTGAGCCACCGCGTCCAGCGCCCTGGCGGTTTAAATCA	420	


```

; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSEQ for Windows Version 3.0

```

```

; SEQ ID NO 7
; LENGTH: 72604
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified base
; LOCATION: all n positions
; OTHER INFORMATION: n=a, c, g, or t
US-09-657-474-7

Query Match      3.9%; Score 46; DB 4; Length 72604;
Best Local Similarity 100.0%; Pred. No. 6.8e-08;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 183 GATCTGGCTCACTGCAACCTCTGCTCCAGGTTCAAGCAATTCT 228
|||||
Db 46184 GATCTGGCTCACTGCAACCTCTGCTCCAGGTTCAAGCAATTCT 46229

RESULT 7
US-09-038-832-1
; Sequence 1, Application US/09038832
; Patent No. 6146845
; GENERAL INFORMATION:
; APPLICANT: KIKLY, KRISTINE
; APPLICANT: BRICKSON-MILLER, CONNIE
; TITLE OF INVENTION: Sialoadhesin Family Member-2
; TITLE OF INVENTION: (SAF-2)
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RATNER & PRESTIA
; STREET: P.O. BOX 980
; CITY: VALLEY FORGE
; STATE: PA
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/038,832
; FILING DATE: 11-MAR-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/041,886
; FILING DATE: 02-APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: PRESTIA, PAUL F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: GH-50018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0701
; TELEX: 846169
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2900 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-09-038-832-1

Query Match      3.4%; Score 40; DB 3; Length 2900;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 189 GGCTCACTGCAACCTCTGCTCCAGGTTCAAGCAATTCT 228
|||||
Db 1881 GGCTCACTGCAACCTCTGCTCCAGGTTCAAGCAATTCT 1920
```

```

RESULT 8
US-09-103-875-2/c
; Sequence 2, Application US/09103875A
; Patent No. 6221849
; GENERAL INFORMATION:
; APPLICANT: Szyf, Moshe
; APPLICANT: Bigey, Pascal
; APPLICANT: Ramchandani, Shyam
; TITLE OF INVENTION: DNA METHYLTRANSFERASE GENOMIC SEQUENCES AND ANTISENSE
; TITLE OF INVENTION: OLIGONUCLEOTIDES
; FILE REFERENCE: 106101.194
; CURRENT APPLICATION NUMBER: US/09/103,875A
; CURRENT FILING DATE: 1998-06-24
; EARLIER APPLICATION NUMBER: 60/069,865
; EARLIER FILING DATE: 1997-12-17
; EARLIER APPLICATION NUMBER: 08/866,340
; EARLIER FILING DATE: 1997-05-30
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 4136
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-103-875-2

Query Match      3.4%; Score 40; DB 4; Length 4136;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 183 GATCTGGCTCACTGCAACCTCTGCTCCAGGTTCAAGC 222
|||||
Db 2510 GATCTGGCTCACTGCAACCTCTGCTCCAGGTTCAAGC 2471

RESULT 9
US-08-306-691B-19
; Sequence 19, Application US/08306691B
; Patent No. 5734039
; GENERAL INFORMATION:
; APPLICANT: Calabretta, Bruno
; APPLICANT: Skorski, Tomasz
; TITLE OF INVENTION: ANTISENSE
; TITLE OF INVENTION: OLIGONUCLEOTIDES TARGETING COOPERATING ONCOGENES
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seidel, Gonda, Javorigna & Monaco, P.C.
; STREET: Two Penn Center, Suite 1800
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/306,691B
; FILING DATE: September 15, 1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 8321-8
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; TELEX: No. 5734039e
; INFORMATION FOR SEQ ID NO: 19:
```

```

SEQUENCE CHARACTERISTICS:
LENGTH: 35100 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-306-691B-19

Query Match          3.4%; Score 40; DB 1; Length 35100;
Best Local Similarity 100.0%; Pred. No. 8.4e-06;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 189 GGTCACTGCAACCTCTGCTCCAGGTTCAAGCAATTCT 228
DB 28051 GGTCACTGCAACCTCTGCTCCAGGTTCAAGCAATTCT 28090

RESULT 10
PCT-US93-06251-19
; Sequence 19, Application PC/TUS9306251
; GENERAL INFORMATION:
; APPLICANT: Wickstrom, Eric and wife, Jason P.
; TITLE OF INVENTION: Trivalent Synthesis of Oligonucleotides Containing
; TITLE OF INVENTION: Stereospecific Alkylphosphonates and Arylphosphonates
; NUMBER OF SEQUENCES: 93
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: NY
; COUNTRY: USA
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/06251
; FILING DATE: 19930630
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 8586
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 516-742-4343
; TELEFAX: 516-742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35100 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
PCT-US93-06251-19

Query Match          3.4%; Score 40; DB 5; Length 35100;
Best Local Similarity 100.0%; Pred. No. 8.4e-06;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 189 GGTCACTGCAACCTCTGCTCCAGGTTCAAGCAATTCT 228
DB 28051 GGTCACTGCAACCTCTGCTCCAGGTTCAAGCAATTCT 28090

RESULT 11
US-09-257-179-32/c
; Sequence 32, Application US/09257179
; Patent No. 6410709
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 29 Human Secreted Proteins

```

```

FILE REFERENCE: PZ015PI
CURRENT APPLICATION NUMBER: US/09/257,179
CURRENT FILING DATE: 1999-02-25
EARLIER APPLICATION NUMBER: PCT/US98/17709
EARLIER FILING DATE: 1998-08-27
EARLIER APPLICATION NUMBER: 60/056,270
EARLIER FILING DATE: 1997-08-29
EARLIER APPLICATION NUMBER: 60/056,271
EARLIER FILING DATE: 1997-08-29
EARLIER APPLICATION NUMBER: 60/056,247
EARLIER FILING DATE: 1997-08-29
EARLIER APPLICATION NUMBER: 60/056,073
EARLIER FILING DATE: 1997-08-29
NUMBER OF SEQ ID NOS: 128
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 32
LENGTH: 1014
TYPE: DNA
ORGANISM: Homo sapiens
US-09-257-179-32

Query Match          3.3%; Score 39; DB 4; Length 1014;
Best Local Similarity 100.0%; Pred. No. 3.6e-05;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 190 GGTCACTGCAACCTCTGCTCCAGGTTCAAGCAATTCT 228
DB 940 GGTCACTGCAACCTCTGCTCCAGGTTCAAGCAATTCT 902

RESULT 12
US-08-687-080-76/c
; Sequence 76, Application US/08687080
; Patent No. 5965427
; GENERAL INFORMATION:
; APPLICANT: Gregory Dolganov
; TITLE OF INVENTION: Human RAD50 Gene and Methods of Use Thereof
; NUMBER OF SEQUENCES: 175
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/687,080
; FILING DATE: 17-JUL-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/592,126
; FILING DATE: 26-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Sholtz, Charles K.
; REGISTRATION NUMBER: 38,615
; REFERENCE/DOCKET NUMBER: 4600-0111.30
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 76:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1386 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO

```


ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: INTRON 9 OF RAD50 GENOMIC SEQUENCE
US-08-687-080-76

Query Match 3.3%; Score 39; DB 2; Length 1386;
Best Local Similarity 100.0%; Pred. No. 3.4e-05;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 184 ACTTGCTCACTGCAACCTCTGCTCCAGGTTCAAGC 222
|||||
DB 586 ACTTGCTCACTGCAACCTCTGCTCCAGGTTCAAGC 548

RESULT 13

US-08-938-669A-1
; Sequence 1, Application US/08938669A
; Patent No. 6171788
; GENERAL INFORMATION:
; APPLICANT: Nguyen, Thai D.
; TITLE OF INVENTION: METHODS FOR THE DIAGNOSIS,
; TITLE OF INVENTION: PROGNOSIS AND TREATMENT OF GLAUCOMA AND
; TITLE OF INVENTION: RELATED DISEASES
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howrey & Simon
; STREET: 1299 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20004-2402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/938,669A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/791,154
; FILING DATE: 28-JAN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mendelson, Elliot
; REGISTRATION NUMBER: P-42,878
; REFERENCE/DOCKET NUMBER: 07425-0034
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202 383-6857
; TELEFAX: 202 383-6610
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5300 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-938-669A-1

Query Match 3.3%; Score 39; DB 4; Length 5300;
Best Local Similarity 100.0%; Pred. No. 2.6e-05;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 190 GCTCACTGCAACCTCTGCTCCAGGTTCAAGCAATTCT 228
|||||
DB 1327 GCTCACTGCAACCTCTGCTCCAGGTTCAAGCAATTCT 1365

RESULT 14

US-08-938-669A-2
; Sequence 2, Application US/08938669A
; Patent No. 6171788
; GENERAL INFORMATION:

APPLICANT: Nguyen, Thai D.
APPLICANT: Polansky, Jon R.
TITLE OF INVENTION: METHODS FOR THE DIAGNOSIS,
TITLE OF INVENTION: PROGNOSIS AND TREATMENT OF GLAUCOMA AND
TITLE OF INVENTION: RELATED DISEASES
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Howrey & Simon
STREET: 1299 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20004-2402
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/938,669A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/791,154
FILING DATE: 28-JAN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Mendelson, Elliot
REGISTRATION NUMBER: P-42,878
REFERENCE/DOCKET NUMBER: 07425-0034
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202 383-6857
TELEFAX: 202 383-6610
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 5304 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-938-669A-2

Query Match 3.3%; Score 39; DB 4; Length 5304;
Best Local Similarity 100.0%; Pred. No. 2.6e-05;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 190 GCTCACTGCAACCTCTGCTCCAGGTTCAAGCAATTCT 228
|||||
DB 1327 GCTCACTGCAACCTCTGCTCCAGGTTCAAGCAATTCT 1365

RESULT 15

US-08-938-669A-3
; Sequence 3, Application US/08938669A
; Patent No. 6171788
; GENERAL INFORMATION:
; APPLICANT: Nguyen, Thai D.
; APPLICANT: Polansky, Jon R.
; TITLE OF INVENTION: METHODS FOR THE DIAGNOSIS,
; TITLE OF INVENTION: PROGNOSIS AND TREATMENT OF GLAUCOMA AND
; TITLE OF INVENTION: RELATED DISEASES
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howrey & Simon
; STREET: 1299 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20004-2402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/938,669A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/791,154
; FILING DATE: 28-JAN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mendelson, Elliot
; REGISTRATION NUMBER: P-42,878
; REFERENCE/DOCKET NUMBER: 07425-0034
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202 383-6857
; TELEFAX: 202 383-6610
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6169 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-08-938-669A-3
Query Match 3.3%; Score 39; DB 4; Length 6169;
Best Local Similarity 100.0%; Pred. No. 2.5e-05;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 190 GCTCACTGCAACCTCTGCTCCAGGTTCAAGCAATTCT 228
|||||
Db 1327 GCTCACTGCAACCTCTGCTCCAGGTTCAAGCAATTCT 1365

```

Search completed: November 29, 2002, 03:34:35
Job time : 288.387 secs

GenCore version 5.1.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

DM nucleic - nucleic search, using sw model

Run on: November 28, 2002, 22:40:30 ; Search time 235.395 Seconds
(without alignments)
11355.910 Million cell updates/sec

Title: US-09-963-333-7
Perfect score: 1187
Sequence: 1 gatcgccactgcactcca.....gagggagggcggtgggga 1187

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 2185239 seqs, 1125999159 residues

Word size : 15

Total number of hits satisfying chosen parameters: 263288

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Result No.	Score	Query % Match	Length	DB ID	Description
1	658	55.4	18596	22	AAF31109
2	658	55.4	18596	22	AAF31109
3	658	55.4	18596	24	ABK43334
4	658	55.4	18596	24	ABK43334
5	658	55.4	18596	24	ABK43334
6	658	55.4	18596	24	ABK43334
7	658	55.4	18596	24	ABK43334
8	297	25.0	1536	24	ABK43335
9	262	22.1	3298	24	AAK94945

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

C 10	205	17.3	1539	23	AAK94960	DNA encoding novel
C 11	165	13.9	1131	24	ABK43330	Human GNHK genomic
C 12	165	13.9	1161	24	ABK43288	Human cDNA encoding
C 13	108	9.1	535	24	ABL38559	Human colon tumour
C 14	60	5.2	7428	22	ABA07164	Human pancreatic c
C 15	60	5.1	566	20	AAK24270	Human thymidylate
C 16	59	5.0	22473	22	AAK71400	Human immune/haema
C 17	58	4.9	3702	22	AAK72556	Human immune/haema
C 18	58	4.9	10867	22	AAK37605	Human musculoskele
C 19	58	4.9	10867	22	AAK36613	Human cardiovascular
C 20	58	4.9	10867	22	AAK66345	Human immune/haema
C 21	58	4.9	10907	22	AAK37603	Human musculoskele
C 22	58	4.9	10907	22	AAK36612	Human cardiovascular
C 23	58	4.9	10907	22	AAK66344	Human immune/haema
C 24	58	4.9	12718	22	AAK37607	Human musculoskele
C 25	58	4.9	12718	22	AAK66346	Human immune/haema
C 26	58	4.9	23989	22	AAK72555	Human immune/haema
C 27	56	4.7	913	22	AAI94312	Human neuroblastom
C 28	56	4.7	5253	22	AAK9859	Human digestive sy
C 29	56	4.7	5257	22	AAK9860	Human digestive sy
C 30	56	4.7	143306	24	ABK49586	Human transporter
C 31	55	4.6	10894	22	AAK36611	Human cardiovascular
C 32	55	4.6	10894	22	AAK66343	Human immune/haema
C 33	54	4.5	10012	24	ABL55889	Human small induci
C 34	54	4.5	58837	24	ABK52612	Human Claspino geno
C 35	50	4.2	38358	22	AAK73555	Human immune/haema
C 36	48	4.0	32170	22	AAK28674	Genomic sequence #
C 37	48	4.0	36501	22	AAK64829	Human immune/haema
C 38	47	4.0	1110	19	AAV29343	Calcium ion channe
C 39	47	4.0	12278	22	ABA14585	Human nervous syst
C 40	47	4.0	12278	22	ABA14664	Human nervous syst
C 41	46	3.9	111	22	AAK06986	Human reproductive
C 42	46	3.9	115	22	AAK30875	Human digestive sy
C 43	46	3.9	255	21	AAK05034	Human secreted pro
C 44	46	3.9	429	22	AAK12261	Human breast cance
C 45	46	3.9	490	22	AAK77714	Human immune/haema

ALIGNMENTS

RESULT 1

AAF31109
ID AAF31109 standard; cDNA; 18596 BP.

XX AAF31109;

XX AAF31109;

DT 27-APR-2001 (first entry)

XX Thymidylate synthase coding sequence.

XX Analyte-binding enzyme; analyte analysis; ss.

XX Homo sapiens.

XX WO200102600-A2.

XX 11-JAN-2001.

XX 30-JUN-2000; 2000WO-US18057.

XX 06-JUL-1999; 99US-0347878.

XX 06-DEC-1999; 99US-0457205.

XX (GEAT) GEN ATOMICS.

XX Yuan C;

XX WPI; 2001-071583/08.

PT Assaying method, useful for prognosis and diagnosis of disease,
PT comprises contacting sample with a mutant analyte-binding enzyme and
PT detecting binding -

XX Disclosure; Page -; 187bp; English.

XX The present invention relates to a method for assaying an analyte in a sample comprising: contacting the sample with a mutant analyte-binding enzyme which has binding affinity for the analyte or an immediate activity, and detecting resulting binding. The method is useful in monitoring biological systems/processes, or prognosis/diagnosis of disease caused by imbalances of the analytes. The present sequence is a coding sequence used in the present invention.

CC Note: the present sequence is not shown in the specification, but was from Genbank, using information given in the specification.

XX

SQ Sequence 18596 BP; 4521 A; 3991 C; 4479 G; 5605 T; 0 other;

Query Match 55.4%; Score 658; DS 22; Length 18596;

Best Local Similarity 99.2%; Pred. No. 6.3e-195;

Matches 1178; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

QY 1 GATCGCGCACTGACCTCCAGCTGGGTGAGAGAGCAGACTCTGTCTCAAAAAA 60

Dd 78 GATCGCGCACTGACCTCCAGCTGGGTGAGAGAGCAGACTCTGTCTCAAAAAA 137

QY 61 AAAAAGACCGCCAGGGTCAAAACAAAACCTCGGAAAGCCCTGCGGCTCTTTT 120

Dd 138 AAAAGACCGCCAGGGTCAAAACAAAACCTCGGAAAGCCCTGCGGCTCTTTT 197

QY 121 TTTTCTTTTCTTTTCTTTTGGGACAGTCTGTCTGTGCGCCAGGCTGAGTACAA 180

Dd 198 TTTTCTTTTCTTTTCTTTTGGGACAGTCTGTCTGTGCGCCAGGCTGAGTACAA 257

QY 181 CGGATCTGGCTCTACTGCACTCTGCTCCAGCTTCAAGCAATCTCTGCTCAGCC 240

Dd 258 CGGATCTGGCTCTACTGCACTCTGCTCCAGCTTCAAGCAATCTCTGCTCAGCC 317

QY 241 TCCCAAGTAGCCACCGCCAGCTAAATTTTGTATTTTAGTAGAGCGGGTTTCA 300

Dd 318 TCCCAAGTAGCCACCGCCAGCTAAATTTTGTATTTTAGTAGAGCGGGTTTCA 377

QY 301 CATGTGTCCAGGCTGTCTGAACTCTGACCTCAGGTGATCCACCGCTCGCC 360

Dd 378 CATGTGTCCAGGCTGTCTGAACTCTGACCTCAGGTGATCCACCGCTCGCC 437

QY 361 CAAAGTACTAGGATTAAGCGCTGAGCCACCGGCTCAGCCCTGCGGTTTAA 420

Dd 438 CAAAGTACTAGGATTAAGCGCTGAGCCACCGGCTCAGCCCTGCGGTTTAA 497

QY 421 AGTAGAAAGCTGCAATTAAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480

Dd 498 AGTAGAAAGCTGCAATTAAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 556

QY 481 AAATGCAATCTTATTAAGTGTAGGAAACAGATCTCAAAACAGCAGTCTTT 540

Dd 557 AAATGCAATCTTATTAAGTGTAGGAAACAGATCTCAAAACAGCAGTCTTT 616

QY 541 ACCGAGGAAACGCTGGAACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600

Dd 617 ACCGAGGAAACGCTGGAACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 676

QY 601 TTCCCAAGGCGCAGTCTTCCCGCCAGCCAGCTGCTGCTGCTGCTGCTGCTGCT 660

Dd 677 TTCCCAAGGCGCAGTCTTCCCGCCAGCCAGCTGCTGCTGCTGCTGCTGCTGCT 736

QY 661 AAGACTCTCAGCTGCGCTTGGCTCCGTTCTGTGCTGCTGCTGCTGCTGCTGCT 720

Dd 737 AAGACTCTCAGCTGCGCTTGGCTCCGTTCTGTGCTGCTGCTGCTGCTGCTGCT 796

QY 721 CCCTTGGCGCAGCTCTTAGAGCGGGGCGCGCCAGCCCGCGGAGGAGAGGCG 780

Dd 797 CCCTTGGCGCAGCTCTTAGAGCGGGGCGCGCCAGCCCGCGGAGGAGAGGCG 856

QY 781 GAGCGGCGGACCGCGCGGGAAGAGCGCGCGGAGGGGTCTGTGCCACCGCGCACTTGG 840

Dd 857 GAGCGCGGACCGCGCGGGGAAAGGCGCGCGGAGGGGTCTGTGCCACCGCGCACTTGG 916

QY 841 CTTGCTCTCTGCTCCCGCGCGCCACCTTGTGCTGCTTCCGTCCCGCGCGGCACTTGTGCTG 900

Dd 917 CTTGCTCTCTGCTCCCGCGCGCCACCTTGTGCTGCTTCCGTCCCGCGCGGCACTTGTGCTG 976

QY 901 CTTGCTCTCTGCTCCCGCGCGCCACCTTGTGCTGCTTCCGTCCCGCGCGGCACTTGTGCTG 960

Dd 977 CTTGCTCTCTGCTCCCGCGCGCCACCTTGTGCTGCTTCCGTCCCGCGCGGCACTTGTGCTG 1036

QY 961 TTGCTCTCTGCTCCCGCGCGCCACCTTGTGCTGCTTCCGTCCCGCGCGGCACTTGTGCTG 1020

Dd 1037 TTGCTCTCTGCTCCCGCGCGCCACCTTGTGCTGCTTCCGTCCCGCGCGGCACTTGTGCTG 1096

QY 1021 TACCTTGGGCGCAGATCCAAACATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080

Dd 1097 TACCTTGGGCGCAGATCCAAACATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1156

QY 1081 ACCGCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140

Dd 1157 ACCGCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1216

QY 1141 GCGCTCTGCGGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1187

Dd 1217 GCGCTCTGCGGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1263

RESULT 2

AAC91215

ID AAC91215 standard; DNA; 18596 BP.

XX AAC91215;

XX 20-MAR-2001 (first entry)

XX Human thymidylate synthase gene SEQ ID NO: 11.

XX Human; schizophrenia; developmental disorder; spina bifida cystica;

KW Tourette's syndrome; bipolar illness; autism; conduct disorder;

KW attention deficit disorder; obsessive compulsive disorder;

KW chronic multiple tic syndrome; learning disorder; polymorphism; ds.

XX Homo sapiens.

XX WC2000071754-A1.

XX 30-NOV-2000.

XX 24-MAY-2000; 2000WO-US14354.

XX 25-MAY-1999; 99US-0318448.

XX (UYNE-) UNIV NEW JERSEY MEDICINE & DENTISTRY.

XX Johnson WG, Stenroos ES;

XX WPI; 2001-025174/03.

XX Diagnosing a developmental disorder, e.g. schizophrenia, by forming

PT datasets (DS) of genetic (e.g. genotypes of folate metabolism alleles)

PT and environmental variables affecting an individual and then comparing

PT these DS with reference DS -

XX Disclosure; Page 125-131; 156pp; English.

XX The present invention provides a novel method of estimating the

CC susceptibility of an individual to a developmental disorder using genetic

CC and environmental variables. The method can be used in the diagnosis,

CC prevention and treatment of disorders such as schizophrenia, spina bifida

CC cystica, Tourette's syndrome, bipolar illness, autism, conduct disorders,

CC attention deficit disorder, obsessive compulsive disorder, chronic

CC multiple tic syndrome and learning disorders such as dyslexia.

QY 1 GATCCGCCACTGCTCAGCTCGAGCTGGTGGAGAGCGAGACTCTCTCTCAAAAAA 60
Db |||||
QY 78 GATCGGCCACTGCTCAGCTCGAGCTGGTGGAGAGCGAGACTCTCTCTCAAAAAA 137
Db |||||
QY 61 AAAAGACCGCCAGGCTCAACAAAAAACCCTCGAAAAACCTCGGGTCTTTT 120
Db |||||
QY 138 AAAAGACCGCCAGGCTCAACAAAAAACCCTCGAAAAACCTCGGGTCTTTT 197
Db |||||
QY 121 TTTTCTTTTCTTTTCTTTTGGGACAGTCTTCTCTGCGCCAGGCTGGAGTACATG 180
Db |||||
QY 198 TTTTCTTTTCTTTTCTTTTGGGACAGTCTTCTCTGCGCCAGGCTGGAGTACATG 257
Db |||||
QY 181 CGGATCTGGCTCAGTCAACACCTCTGCTCCAGGTTCAAGCAATCTTCTGCTCAGCC 240
Db |||||
QY 258 CGGATCTGGCTCAGTCAACACCTCTGCTCCAGGTTCAAGCAATCTTCTGCTCAGCC 317
Db |||||
QY 241 TCCCAGTAGCCACCGCCAGCTAAATTTTGTANTTTTGTAGTAGAGCGGGTTTCA 300
Db |||||
QY 318 TCCCAGTAGCCACCGCCAGCTAAATTTTGTANTTTTGTAGTAGAGCGGGTTTCA 377
Db |||||
QY 301 CATGTCTCCAGGCTGGTCTGAGAACCTCTGACCTCAGTGTATCAACCGCTCGGCC 360
Db |||||
QY 378 CATGTCTCCAGGCTGGTCTGAGAACCTCTGACCTCAGTGTATCAACCGCTCGGCC 437
Db |||||
QY 361 CAAAGTACTAGGATTAAGCGGTGAGCAACCGCTCAGCGCTCTGGCGTTTAAATCA 420
Db |||||
QY 438 CAAAGTACTAGGATTAAGCGGTGAGCAACCGCTCAGCGCTCTGGCGTTTAAATCA 497
Db |||||
QY 421 AGTAGAAAGCTGATTAACCACTTCTGCTTGTGTTGTTTCAAGTGAAGAGAAATGG 480
Db |||||
QY 498 AGTAGAAAGCTGATTAACCACTTCTGCTTGTGTTGTTTCAAGTGAAGAGAAATGG 556
Db |||||
QY 481 AAATGCAATCTTATTTAGTGTAGGAACAGATCTCAACAGCACTTTTGTGACAG 540
Db |||||
QY 557 AAATGCAATCTTATTTAGTGTAGGAACAGATCTCAACAGCACTTTTGTGACAG 616
Db |||||
QY 541 ACGCAGAAAGCTGGGAACTGTGCTGTGCTTAGAAGAGCGCGTCCGACCAAG 600
Db |||||
QY 617 ACGCAGAAAGCTGGGAACTGTGCTGTGCTTAGAAGAGCGCGTCCGACCAAG 676
Db |||||
QY 601 TTCCCAAGAGCGGAGTCTTCCGCGCCAGCACTGCTGCTTCCAGGTTCCCGGTTCT 660
Db |||||
QY 677 TTCCCAAGAGCGGAGTCTTCCGCGCCAGCACTGCTGCTTCCAGGTTCCCGGTTCT 736
Db |||||
QY 661 AAGACTCTCAGTGTGGCCCTGGGCTCGGTCTGTGTCACACCGCTGCTCGCTTTC 720
Db |||||
QY 737 AAGACTCTCAGTGTGGCCCTGGGCTCGGTCTGTGTCACACCGCTGCTCGCTTTC 796
Db |||||
QY 721 CCCTCGCGCAGCTCTTAGAGCGGGGCGCGCGAGCGCGCGAGGAGAGGCG 780
Db |||||
QY 797 CCCTCGCGCAGCTCTTAGAGCGGGGCGCGCGAGCGCGCGAGGAGAGGCG 856
Db |||||
QY 781 GAGCGGAGCGCGCGGAGGAGGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
Db |||||
QY 857 GAGCGGAGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 916
Db |||||
QY 841 CTGCTCTCTGCTCCGCGCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
Db |||||
QY 917 CTGCTCTCTGCTCCGCGCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 976
Db |||||
QY 901 CTGCTCTCTGCTCCGCGCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
Db |||||
QY 977 CTGCTCTCTGCTCCGCGCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1036
Db |||||
QY 961 TTGCTCTCTGCTCCGCGCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
Db |||||
QY 1037 TTGCTCTCTGCTCCGCGCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1096
Db |||||
QY 1021 TACTGGGGAGATTCACACATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
Db |||||
QY 1097 TACTGGGGAGATTCACACATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1156
Db |||||

QY 1081 ACGGCACCTGTCTGCTATTCGGCATGCGAGCGCTACGCTGACAGTGCAGCGCG 1140
Db |||||
QY 1157 ACGGCACCTGTCTGCTATTCGGCATGCGAGCGCTACGCTGACAGTGCAGCGCG 1216
Db |||||
QY 1141 GGCCCTCGCGGAGCGGCTGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1187
Db |||||
QY 1217 GGCCCTCGCGGAGCGGCTGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1263
Db |||||

RESULT 4

ABK43334
ID ABK43334 standard; DNA; 18596 BP.
XX
AC ABK43334;
XX
DT 05-JUN-2002 (first entry)
XX
DE Human Thymidylate synthase gene sequence.

XX HKNG1; ds; gene; chromosome 18p; bipolar affective disorder; BAD;
XX severe bipolar affective (mood) disorder; BP-I; schizophrenia; TS;
XX Hong Kong new gene 1; antimanic; antidepressant; neuroleptic;
XX Thymidylate synthase.

XX Homo sapiens.

XX WC0200210366-A2.

XX PD 07-FEB-2002.

XX 02-AUG-2001; 2001WO-US24417.

XX 02-AUG-2000; 2000US-0631275.

XX 28-NOV-2000; 2000US-0722544.

XX (MILL-) MILLENNIUM PHARM INC.
XX (REGC) UNIV CALIFORNIA.

XX Chen H, Freimer NB, Novak T;

XX WPI; 2002-195962/25.

XX New nucleic acid molecule Hong Kong New Gene 1 (HKNG1), useful for
XX screening for molecules which modulate HKNG1 expression for the
XX treatment of bipolar disorder and schizophrenia -

XX Disclosure; Fig 44; 367pp; English.

XX The invention relates to an isolated nucleic acid molecule comprising a
XX nucleotide sequence that encodes a Hong Kong New Gene (HKNG) 1 gene
XX product. The human gene for HKNG1 is located on chromosome 18p in
XX an area associated with bipolar affective disorder, BAD. Also
XX included are an expression vector comprising the nucleic acid, a method
XX of host cell expressing the nucleic acid, and identifying an individual (at
XX risk of) having HKNG1-mediated disorder comprising detecting the presence
XX or absence of a polymorphism that correlates with an HKNG1 allele
XX associated with the disorder, where the presence of the polymorphism
XX indicates that the individual (is at risk of) having HKNG1-mediated
XX disorder. A (small molecule) compound which modulates (inhibits or
XX potentiates) expression of a HKNG1 gene or gene product in a human
XX individual is useful for the treatment of a HKNG1-mediated disorder
XX such as bipolar affective disorder (BAD), severe bipolar affective (mood)
XX disorder (BP-I) and schizophrenia. The present sequence is a
XX thymidylate synthase, TS, genomic DNA sequence. The gene for TS
XX overlaps that of HKNG1 and therefore TS may also be involved in the
XX diseases listed above.

XX Sequence 18596 BP; 4521 A; 3991 C; 4479 G; 5605 T; 0 other;

Query Match 55.4%; Score 658; DB 24; Length 18596;

Best Local Similarity 99.2%; Pred. No. 63e-195;

Matches 1178; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

18-SRP-2000; 2000US-233617P.
 20-SRP-2000; 2000US-234009P.
 20-SRP-2000; 2000US-234034P.
 20-SRP-2000; 2000US-234052P.
 22-SRP-2000; 2000US-234509P.
 22-SRP-2000; 2000US-234567P.
 22-SRP-2000; 2000US-234923P.
 25-SRP-2000; 2000US-234924P.
 25-SRP-2000; 2000US-235077P.
 25-SRP-2000; 2000US-235082P.
 25-SRP-2000; 2000US-235134P.
 25-SRP-2000; 2000US-235280P.
 26-SRP-2000; 2000US-235637P.
 26-SRP-2000; 2000US-235638P.
 27-SRP-2000; 2000US-235720P.
 27-SRP-2000; 2000US-235840P.
 27-SRP-2000; 2000US-235863P.
 28-SRP-2000; 2000US-236028P.
 28-SRP-2000; 2000US-236032P.
 28-SRP-2000; 2000US-236033P.
 28-SRP-2000; 2000US-236034P.
 28-SRP-2000; 2000US-236109P.
 28-SRP-2000; 2000US-236111P.
 29-SRP-2000; 2000US-236842P.
 29-SRP-2000; 2000US-236891P.
 02-OCT-2000; 2000US-237172P.
 02-OCT-2000; 2000US-237173P.
 02-OCT-2000; 2000US-237294P.
 02-OCT-2000; 2000US-237295P.
 02-OCT-2000; 2000US-237316P.
 03-OCT-2000; 2000US-237425P.
 03-OCT-2000; 2000US-237598P.
 03-OCT-2000; 2000US-237604P.
 03-OCT-2000; 2000US-237606P.
 03-OCT-2000; 2000US-237608P.
 01-NOV-2000; 2000US-244867P.
 01-NOV-2000; 2000US-245084P.
 (AVAL-) AVALON PHARM.
 Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
 Soppet DR, Weaver Z;
 WPI; 2002-188264/24.
 Screening for anti-neoplastic agent involves exposing cells to a
 chemical agent to be tested for anti-neoplastic activity, and
 determining a change in expression of a gene of a signature gene set -
 Claim 1; SEQ ID 1415; 44pp; English.
 The present invention describes a method (M1) for screening for an
 anti-neoplastic agent. The method involves exposing cells to a chemical
 agent to be tested for anti-neoplastic activity, determining a change in
 expression of at least one gene (I) of a signature gene set, where (I)
 comprises a sequence (S) selected from 847 sequences (given in ABL6164
 to ABL70110), or is at least 95% identical to (S), where a change in
 expression is indicative of anti-neoplastic activity. (I) has cytostatic
 activity and can be used in gene therapy. M1 can be used for screening
 an anti-neoplastic agent, and can be used for producing a product which
 is the data collected with respect to the anti-neoplastic agent as a
 result of M1, and the data is sufficient to convey the chemical
 structure and/or properties of the agent. M1 can be used in the
 treatment of cancer such as colon, breast, stomach, lung, thyroid,
 oesophageal, ovarian, kidney, prostate or pancreatic cancer,
 adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
 cc
 infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
 carcinoma, papillary carcinoma and wilm's tumour.
 XX
 Sequence 18596 BP; 4521 A; 3991 C; 4479 G; 5605 T; 0 other;
 SQ

Query Match 55.4%; Score 658; DB 24; Length 18596;
 Best Local Similarity 99.2%; Pred. No. 6.3e-195;
 Matches 1178; Conservative 0; Mismatches 8; Indels 1; Gaps 1;
 Qy 1 GATCGCGCACTGCACCTCCAGCTGGGTGAGAGAGCGAGACTCTGTCTCAAAAAA 60
 Db 78 GATCGCGCACTGCACCTCCAGCTGGGTGAGAGAGCGAGACTCTGTCTCAAAAAA 137
 Qy 61 AAAAGAGCGCGCAGGGCTCAAAACAAAAACCTCGGAAAAGCCCTGGGGTCTTTTTTT 120
 Db 138 AAAAGAGCGCGCAGGGCTCAAAACAAAAACCTCGGAAAAGCCCTGGGGTCTTTTTTT 197
 Qy 121 TTTTTTTTTTTTTTTTTTGGGACAGTCTGTCTGTGGCCAGGCTGGAGTACAATGGT 180
 Db 198 TTTTTTTTTTTTTTTTTTGGGACAGTCTGTCTGTGGCCAGGCTGGAGTACAATGGT 257
 Qy 181 CGGATCTTGGCTCACTGCAACCTCTGCTCCAGGTTCAAGCAATTTCTTCTGCTCAGCC 240
 Db 258 CGGATCTTGGCTCACTGCAACCTCTGCTCCAGGTTCAAGCAATTTCTTCTGCTCAGCC 317
 Qy 241 TCCAGTAGCCACACCGCCAGCTAAATTTTGTANTTTTAGTAGAGCGGGGTTTCAC 300
 Db 318 TCCAGTAGCCACACCGCCAGCTAAATTTTGTACTTTTAGTAGAGCGGGGTTTCAC 377
 Qy 301 CATGTTGTCAGGCTGTCTGTAACCTCTGACCTCAGGTGATCCACCCGCTCGGCCCCC 360
 Db 378 CATGTTGTCAGGCTGTCTGTAACCTCTGACCTCAGGTGATCCACCCGCTCGGCCCCC 437
 Qy 361 CAAAGTACTAGGATTACAGCGGTGAGCCACCGCTCCAGCGCTCGCGGTTTTTAATCA 420
 Db 438 CAAAGTACTAGGATTACAGCGGTGAGCCACCGCTCCAGCGCTCGCGGTTTTTAATCA 497
 Qy 421 AGTAGAAAAAGCTGCATTATACCACTTGTCTTNGTTTGNNTTCAGTGAGAACGAAATGG 480
 Db 498 AGTAGAAAAAGCTGCATTATACCACTTGTCTTGGTTTC-TTCAGTGAGAACGAAATGG 556
 Qy 481 AAATGCAATCNCITTTAGTTAGTGTAGAAAACAGATCTCAAAACAGCAGTGTGTTGACAAG 540
 Db 557 AAATGCAATCNCITTTAGTTAGTGTAGAAAACAGATCTCAAAACAGCAGTGTGTTGACAAG 616
 Qy 541 ACCGAGAGAAAACGTGGAACTGTGCTGTGGCTTAGAGAGCGGGTCCGACACGAGCG 600
 Db 617 ACCGAGAGAAAACGTGGAACTGTGCTGTGGCTTAGAGAGCGGGTCCGACACGAGCG 676
 Qy 601 TTCCCAAGAGCGCAGTCTTCCCGCCACCGCAGCTGNTCCAGGTTCCCGGGTTCCT 660
 Db 677 TTCCCAAGAGCGCAGTCTTCCCGCCACCGCAGCTGNTCCAGGTTCCCGGGTTCCT 736
 Qy 661 AAGACTCTCAGCTGTGGCTTCCGCTTCTGTGCCACACCGCTGGCTTCTGCGTTTC 720
 Db 737 AAGACTCTCAGCTGTGGCTTCCGCTTCTGTGCCACACCGCTGGCTTCTGCGTTTC 796
 Qy 721 CCCCTGGCGCAGCTCTCTTAGAGCGGGGCGCCGCGACCCCGCAGCAGGAGAGCGG 780
 Db 797 CCCCTGGCGCAGCTCTCTTAGAGCGGGGCGCCGCGACCCCGCAGCAGGAGAGCGG 856
 Qy 781 GAGCGCGGAGCGCGCGGAAAAGCGCGGAGAGGGGTCTTCCACCGCGCCTCTGG 840
 Db 857 GAGCGCGGAGCGCGCGGAAAAGCGCGGAGAGGGGTCTTCCACCGCGCCTCTGG 916
 Qy 841 CCTGCTCTGCTCGCGCGCAGCTTGGCTTCCGCTTCCCGCGCGCAGCTTGGCTG 900
 Db 917 CCTGCTCTGCTCGCGCGCAGCTTGGCTTCCGCTTCCCGCGCGCAGCTTGGCTG 976
 Qy 901 CCTTCCGTCCTCGCGCGCAGCTTGGCTTCCGCTTCCCGCGCGCAGCTTGGCTG 960
 Db 977 CCTTCCGTCCTCGCGCGCAGCTTGGCTTCCGCTTCCCGCGCGCAGCTTGGCTG 1036
 Qy 961 TTGCCCCCGCGCAGGAGCGGAGCGCGAGCCGCTCGCGCGCAGCGGAGCTGCGAG 1020
 Db 1037 TTGCCCCCGCGCAGGAGCGGAGCGCGAGCCGCTCGCGCGCAGCGGAGCTGCGAG 1096
 Qy 1021 TACCTGGGCGAGATCCAAACATCTCTCCGCTGCGGCTCAGGAAGGACGACGCGACGGCG 1080

02-AUG-2000; 2000US-0631275.
28-NOV-2000; 2000US-0722544.
(MILL-) MILLENNIUM PHARM INC.
(RECG) UNIV CALIFORNIA.
Chen H, Freimer NB, Novak T;
WPI; 2002-195962/25.
P-PSDB; AAU87071.
New nucleic acid molecule Hong Kong New Gene 1 (HKNG1), useful for
screening for molecules which modulate HKNG1 expression for the
treatment of bipolar disorder and schizophrenia -
Disclosure; Fig 45; 367pp; English.
The invention relates to an isolated nucleic acid molecule comprising a
nucleotide sequence that encodes a Hong Kong New Gene (HKNG) 1 gene
product. The human gene for HKNG1 is located on chromosome 18p in
an area associated with bipolar affective disorder, BAD. Also
included are an expression vector comprising the nucleic acid, a
host cell expressing the nucleic acid, an anti-HKNG1 antibody, a method
of identifying modulators of HKNG1, and identifying an individual (at
risk of) having HKNG1-mediated disorder comprising detecting the presence
or absence of a polymorphism that correlates with an HKNG1 allele
associated with the disorder, where the presence of the polymorphism
indicates that the individual (is at risk of) having HKNG1-mediated
disorder. A (small molecule) compound which modulates (inhibits or
potentiates) expression of a HKNG1 gene or gene product in a human
individual is useful for the treatment of a HKNG1-mediated disorder
such as bipolar affective disorder (BAD), severe bipolar affective (mood)
disorder (BP-I) and schizophrenia. The present sequence is the cDNA
encoding thymidylate synthase, TS. The gene for TS
overlaps that of HKNG1 and therefore TS may also be involved in the
diseases listed above.

Sequence 1536 BP; 390 A; 369 C; 399 G; 378 T; 0 other;
Query Match 25.0%; Score 297; DB 24; Length 1536;
Best Local Similarity 100.0%; Pred. No. 3.4e-83;
Matches 297; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 833 CCACATTGGCCTGCTCGTCCCGCGCCGACATTGGCCTGCTCCGCTCCCGCGCCGAC 892
Db 14 CCACATTGGCCTGCTCGTCCCGCGCCGACATTGGCCTGCTCCGCTCCCGCGCCGAC 73
QY 893 TTGCGCTGCTCGTCCCGCGCCGCGCCGACATTGGCCTGCTCCGCTCCCGCGCCGAC 952
Db 74 TTGCGCTGCTCGTCCCGCGCCGCGCCGACATTGGCCTGCTCCGCTCCCGCGCCGAC 133
QY 953 GCGCGCCCTTCCCGCGCCGCGACAGGAGCGGACGCGCGCGCTCCCGCGCCGACGCGG 1012
Db 134 GCGCGCCCTTCCCGCGCCGCGCGACAGGAGCGGACGCGCGCGCTCCCGCGCCGACGCGG 193
QY 1013 AGCTGCGAGTACCTGGGCGAGATCCAAACATCTTCGCTCCGCTCCGCGCTCGAGCTGCGC 1072
Db 194 AGCTGCGAGTACCTGGGCGAGATCCAAACATCTTCGCTCCGCTCCGCGCTCGAGCTGCGC 253
QY 1073 GCACGGGACCGCGACCCCTGTGGTATTTCGGCATGCGAGGCGCGCTACAGCTGAGAG 1129
Db 254 GCACGGGACCGCGACCCCTGTGGTATTTCGGCATGCGAGGCGCGCTACAGCTGAGAG 310

RESULT 9
AAS94945
ID AAS94945 standard; DNA; 3298 BP.
XX AAS94945;
AC
XX
DT 14-FEB-2002 (first entry)
XX
DE Human DNA sequence #200 expressed during foam cell differentiation.


```
QY 1105 ATGCAGGCGGCTACAGCTGAGAG 1129
|||||
Db 1359 ATGCAGGCGGCTACAGCTGAGAG 1335
|||||

RESULT 11
ABK43330/c
ID ABK43330 standard; DNA; 1131 BP.
AC ABK43330;
XX
XX
DT 05-JUN-2002 (first entry)
XX
DE Human GNHK genomic DNA.
XX
XX HKNG1; ds; gene; chromosome 18p; bipolar affective disorder; BAD;
KW severe bipolar affective (mood) disorder; BP-I; schizophrenia;
KW Hong Kong new gene 1; antimanic; antidepressant; neuroleptic.
XX
XX Homo sapiens.
OS
XX WO200210366-A2.
XX
XX 07-FEB-2002.
PD
XX
XX 02-AUG-2001; 2001WO-US24417.
XX
XX 02-AUG-2000; 2000US-0631275.
XX
XX 28-NOV-2000; 2000US-0722544.
XX
XX (MILL-) MILLENNIUM PHARM INC.
PA (REGC ) UNIV CALIFORNIA.
PI
PI Chen H, Freimer NB, Novak T;
XX
XX WPI; 2002-195962/25.
XX
XX New nucleic acid molecule Hong Kong New Gene 1 (HKNG1), useful for
PT screening for molecules which modulate HKNG1 expression for the
PT treatment of bipolar disorder and schizophrenia -
XX
XX Example 16; Fig 30; 367pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule comprising a
CC nucleotide sequence that encodes a Hong Kong New Gene (HKNG) 1 gene
CC product. The human gene for HKNG1 is located on chromosome 18p in
CC an area associated with bipolar affective disorder, BAD. Also
CC included are an expression vector comprising the nucleic acid, a
CC host cell expressing the nucleic acid, an anti-HKNG1 antibody, a method
CC of identifying modulators of HKNG1, and identifying an individual (at
CC risk of) having HKNG1-mediated disorder comprising detecting the presence
CC or absence of a polymorphism that correlates with an HKNG1 allele
CC associated with the disorder, where the presence of the polymorphism
CC indicates that the individual (is at risk of) having HKNG1-mediated
CC disorder. A (small molecule) compound which modulates (inhibits or
CC potentiates) expression of a HKNG1 gene or gene product in a human
CC individual is useful for the treatment of a HKNG1-mediated disorder
CC such as bipolar affective disorder (BAD), severe bipolar affective (mood)
CC disorder (BP-I) and schizophrenia. The present sequence is HKNG1
CC genomic DNA sequence.
XX
XX Sequence 1131 BP; 255 A; 316 C; 304 G; 254 T; 2 other;
SQ
Query Match 13.9%; Score 165; DB 24; Length 1131;
Best Local Similarity 100.0%; Pred. No. 2.1e-42;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1023 CTGGGGGAGATCCACATCTCCGCTGGCGCTCAGGAGACACCGACGGGCAC 1082
|||||
Db 789 CTGGGGGAGATCCACATCTCCGCTGGCGCTCAGGAGACACCGACGGGCAC 730
|||||

QY 1083 CGGACCCCTGTCGGTATTCCGATGACGGCGCTACAGCTGAGAGTGAACCGCGGG 1142
|||||
```

```
Db 729 CGGACCCCTGTCGGTATTCCGATGACGGCGCTACAGCTGAGAGTGAACCGCGGG 670
QY 1143 CCCCTGCGGGACGGGTGCGGGAAGGAGGAGCGCGCTGGGGA 1187
|||||
Db 669 CCCCTGCGGGACGGGTGCGGGAAGGAGGAGCGCGCTGGGGA 625
|||||
```

```
RESULT 12
ABK43288/c
ID ABK43288 standard; cDNA; 1161 BP.
XX
XX ABK43288;
XX
XX 05-JUN-2002 (first entry)
XX
XX Human cDNA encoding GNKH protein.
DE
XX HKNG1; chromosome 18p; bipolar affective disorder; BAD; GNKH1; ss;
KW severe bipolar affective (mood) disorder; BP-I; schizophrenia; gene;
KW Hong Kong new gene 1; antimanic; antidepressant; neuroleptic.
XX
XX Homo sapiens.
OS
XX WO200210366-A2.
XX
XX 07-FEB-2002.
PD
XX
XX 02-AUG-2001; 2001WO-US24417.
XX
XX 02-AUG-2000; 2000US-0631275.
XX
XX 28-NOV-2000; 2000US-0722544.
XX
XX (MILL-) MILLENNIUM PHARM INC.
PA (REGC ) UNIV CALIFORNIA.
PI
PI Chen H, Freimer NB, Novak T;
XX
XX WPI; 2002-195962/25.
XX
XX New nucleic acid molecule Hong Kong New Gene 1 (HKNG1), useful for
PT screening for molecules which modulate HKNG1 expression for the
PT treatment of bipolar disorder and schizophrenia -
XX
XX Example 16; Fig 28; 367pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule comprising a
CC nucleotide sequence that encodes a Hong Kong New Gene (HKNG) 1 gene
CC product. The human gene for HKNG1 is located on chromosome 18p in
CC an area associated with bipolar affective disorder, BAD. Also
CC included are an expression vector comprising the nucleic acid, a
CC host cell expressing the nucleic acid, an anti-HKNG1 antibody, a method
CC of identifying modulators of HKNG1, and identifying an individual (at
CC risk of) having HKNG1-mediated disorder comprising detecting the presence
CC or absence of a polymorphism that correlates with an HKNG1 allele
CC associated with the disorder, where the presence of the polymorphism
CC indicates that the individual (is at risk of) having HKNG1-mediated
CC disorder. A (small molecule) compound which modulates (inhibits or
CC potentiates) expression of a HKNG1 gene or gene product in a human
CC individual is useful for the treatment of a HKNG1-mediated disorder
CC such as bipolar affective disorder (BAD), severe bipolar affective (mood)
CC disorder (BP-I) and schizophrenia. The present sequence is a cDNA
CC encoding the GNKH1 protein, a protein which is encoded on the opposite
CC strand of DNA encoding HKNG1 and may also be involved in the diseases
CC listed above.
XX
XX Sequence 1161 BP; 285 A; 316 C; 304 G; 254 T; 2 other;
SQ
Query Match 13.9%; Score 165; DB 24; Length 1161;
Best Local Similarity 100.0%; Pred. No. 2.1e-42;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1023 CTGGGGGAGATCCACATCTCCGCTGGCGCTCAGGAGACACCGACGGGCAC 1082
|||||
```

```
Db 789 CCTGGGCGAGATCCACACATCTCCGCTGGCGGCTCAGAAAGGACGACCGCACGGGCAC 730
Qy 1083 CGGCACCTGTGCGTATTCGGCANGCAGGCGCGCTACAGCTGAGAGTGACGCCGCGG 1142
Dy 729 CGGCACCTGTGCGTATTCGGCATGCGGCGCGCTACAGCTGAGAGTGACGCCGCGG 670
Qy 1143 CCCTGCGGACCGGTTGGCGGAAGAGGAGGCGCGCTGGGGA 1187
Dy 669 CCCCTGGGACCGGTTGGCGGAAGAGGAGGCGCGCTGGGGA 625

RESULT 13
ID ABL38559 standard; cDNA; 535 BP.
XX ABL38559;
XX AC
XX XX
XX 08-APR-2002 (first entry)
XX DE Human colon tumour antigen polynucleotide SEQ ID NO:2148.
XX KW Human; colon cancer; colon tumour antigen; cytostatic; vaccine;
XX KW colon tumour metastatic antigen; diagnosis; gene; ss.
XX OS Homo sapiens.
XX XX
XX PN WO200196388-A2.
XX PD 20-DEC-2001.
XX PF 08-JUN-2001; 2001WO-US18557.
XX PR 09-JUN-2000; 2000US-210899P.
XX PR 20-FEB-2001; 2001US-270216P.
XX PA (CORI-) CORIXA CORP.
XX PI Jiang Y, Harlocker SL, Secret H;
XX XX WPI; 2002-114514/15.
XX XX Novel isolated colon tumor polynucleotide differentially expressed in
PT colon tumor or colon metastatic tumor and polypeptides encoded by them,
PT useful for inhibiting development of cancer in patient -
XX Claim 1; SEQ ID 2148; 105pp; English.
XX CC ABL36412 to ABL38645 represent human colon tumour antigen cDNA clones (I)
CC which were isolated from human colon tumour and colon metastatic tumour
CC cDNA libraries. (I) have cytostatic activity and can be used in vaccine
CC production. (I) can be used for stimulating and/or expanding T cells
CC specific for a tumour protein on contact with the T cells. They are also
CC useful for inhibiting the development of cancer in a patient. (I) can be
CC used as probes or primers for nucleic acid hybridisation, for preparing
CC mutant species primers, or primers for use in genetic constructions. (I)
CC can be used in the diagnosis of a colon tumour.
XX SQ Sequence 535 BP; 135 A; 129 C; 151 G; 118 T; 2 other;

Query Match 9.1%; Score 108; DB 24; Length 535;
Best Local Similarity 100.0%; Pred. No. 1e-24; 0; Indels 0; Gaps 0;
Matches 108; Conservative 0; Mismatches 0;

Qy 1022 ACCTGGGCGAGATCCACACATCTCCGCTGGCGGTCAGAAAGGACGACCGCACGGGCAC 1081
Dy 1 ACCTGGGCGAGATCCACACATCTCCGCTGGCGGTCAGAAAGGACGACCGCACGGGCAC 60
Qy 1082 CGGCACCTGTGCGTATTCGGCATGCGGCGCGCTACAGCTGAGAGTGACGCCGCGG 1129
Dy 61 CGGCACCTGTGCGTATTCGGCATGCGGCGCGCTACAGCTGAGAGTGACGCCGCGG 108

RESULT 14
```

```
ABA07164
ID ABA07164 standard; DNA; 7428 BP.
XX AC
XX ABA07164;
XX DT 14-JAN-2002 (first entry)
XX DE Human pancreatic cancer related genomic DNA, SEQ ID NO: 483.
XX KW Human; cytostatic; antidiabetic; antiinflammatory; gastric; osteopathic;
XX KW antihormone; antiulcer; thyroid-active; gene therapy; antisense therapy;
XX KW pancreatic cancer antigen inhibitor; pancreatic cancer; pancreatitis;
XX KW diabetes; endocrine disorder; acromegaly; hyperthyroidism;
XX KW gastrointestinal disorder; Crohn's disease; duodenal ulcer; ds.
XX OS Homo sapiens.
XX XX
XX PN WO200155206-A1.
XX PD 02-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US01353.
XX PR 31-JAN-2000; 2000US-0179065.
XX PR 04-FEB-2000; 2000US-0180628.
XX PR 24-FEB-2000; 2000US-0184664.
XX PR 02-MAR-2000; 2000US-0186350.
XX PR 16-MAR-2000; 2000US-0189874.
XX PR 17-MAR-2000; 2000US-0190076.
XX PR 18-APR-2000; 2000US-0198123.
XX PR 19-MAY-2000; 2000US-0205515.
XX PR 07-JUN-2000; 2000US-0209467.
XX PR 28-JUN-2000; 2000US-0214886.
XX PR 30-JUN-2000; 2000US-0215135.
XX PR 07-JUL-2000; 2000US-0216647.
XX PR 07-JUL-2000; 2000US-0216880.
XX PR 11-JUL-2000; 2000US-0217487.
XX PR 11-JUL-2000; 2000US-0217496.
XX PR 14-JUL-2000; 2000US-0218290.
XX PR 26-JUL-2000; 2000US-0220963.
XX PR 26-JUL-2000; 2000US-0220964.
XX PR 14-AUG-2000; 2000US-0224518.
XX PR 14-AUG-2000; 2000US-0224519.
XX PR 14-AUG-2000; 2000US-0225213.
XX PR 14-AUG-2000; 2000US-0225214.
XX PR 14-AUG-2000; 2000US-0225266.
XX PR 14-AUG-2000; 2000US-0225267.
XX PR 14-AUG-2000; 2000US-0225268.
XX PR 14-AUG-2000; 2000US-0225270.
XX PR 14-AUG-2000; 2000US-0225447.
XX PR 14-AUG-2000; 2000US-0225757.
XX PR 14-AUG-2000; 2000US-0225758.
XX PR 14-AUG-2000; 2000US-0225759.
XX PR 18-AUG-2000; 2000US-0226279.
XX PR 22-AUG-2000; 2000US-0226681.
XX PR 22-AUG-2000; 2000US-0226686.
XX PR 22-AUG-2000; 2000US-0227182.
XX PR 23-AUG-2000; 2000US-0227009.
XX PR 30-AUG-2000; 2000US-0228924.
XX PR 01-SEP-2000; 2000US-0229287.
XX PR 01-SEP-2000; 2000US-0229343.
XX PR 01-SEP-2000; 2000US-0229344.
XX PR 01-SEP-2000; 2000US-0229345.
XX PR 05-SEP-2000; 2000US-0229509.
XX PR 05-SEP-2000; 2000US-0229513.
XX PR 06-SEP-2000; 2000US-0230437.
XX PR 06-SEP-2000; 2000US-0230438.
XX PR 08-SEP-2000; 2000US-0231242.
XX PR 08-SEP-2000; 2000US-0231243.
XX PR 08-SEP-2000; 2000US-0231244.
XX PR 08-SEP-2000; 2000US-0231413.
XX PR 08-SEP-2000; 2000US-0231414.
XX PR 08-SEP-2000; 2000US-0232080.
```


FN WO9915648-A1.
XX
PD 01-APR-1999.
XX
PF 17-SEP-1998; 98WO-GB02820.
XX
XX 06-JUN-1998; 98GB-0012140.
PR 23-SEP-1997; 97GB-0020107.
PR 17-OCT-1997; 97GB-0022012.
XX
PA (ISIS-) ISIS PHARM INC.
PA (ZENE) ZENECA LTD.
XX
PI Dean NM, Koropatnick DJ, Vincent MD;
XX
DR WPI; 1999-254708/21.
XX
PT New antisense oligonucleotides directed against thymidylate synthase
XX
PS Disclosure; Fig 7; 53pp; English.
XX
CC This invention describes novel antisense oligonucleotides targeted to
CC sequences in the 3' end of thymidylate synthase (TS) mRNA. Such
CC oligonucleotides are cytostatic on their own when administered to
CC human tumour cell lines, and also enhance the toxicity of anticancer
CC drugs such as Tomudex administered to those cells. In addition, antisense
CC oligonucleotides targeted to 5' sequences induce TS gene transcription.
CC The antisense oligonucleotides are used in a method for the
CC treatment of cancer (or a method for providing antiproliferative effect)
CC The antisense oligonucleotides are also used in the production of a
CC medicament for the treatment of cancer, either separately or in conjunction
CC with a therapeutic agent such as thymidylate synthase (TS) inhibitors
CC (e.g. Tomudex, Zeneca development compound ZD9331 etc.); cytostatic
CC agents (e.g. antioestrogens, (anti)progestogens, antiandrogens,
CC testosterone inhibitors, anti-invasion agents, growth factor inhibitors,
CC etc.); antiproliferative/antineoplastic agents (e.g. antimetabolites,
CC antitumor antibiotics, alkylating agents, antimitotic agents,
CC topoisomerase inhibitors, etc.), or radiotherapy.
XX
SQ Sequence 566 BP; 156 A; 101 C; 135 G; 174 T; 0 other;

Query Match 5.1%; Score 60; DB 20; Length 566;
Best Local Similarity 100.0%; Pred. No. 6.6e-10;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 900 GCCTCCGTCCTCCCGCCGCGCCATGCTGTGCGCGGCTCGAGCTCCCGCCGCC 959
|||||
Db 1 GCCTCCGTCCTCCCGCCGCGCCATGCTGTGCGCGGCTCGAGCTCCCGCCGCC 60
|||||

Search completed: November 29, 2002, 00:35:19
Job time : 259.395 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

M nucleic - nucleic search, using sw model

Run on: November 28, 2002, 22:48:55 ; Search time 3021.34 Seconds
(without alignments)
11433.695 Million cell updates/sec

Title: US-09-963-333-7

Perfect score: 1187

Sequence: 1 gatcgccactgcactcca.....gagggaggcggtgtgsgga 1187

Scoring table:

OLIGO NUC

Gapop '60.0 , Gapext 60.0

Searched: 2054640 seqs, 14551402878 residues

Word size : 15

Total number of hits satisfying chosen parameters: 471885

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.em.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

15: em.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.mu.*

20: em.om.*

21: em.or.*

22: em.ov.*

23: em.pat.*

24: em.ph.*

25: em.pl.*

26: em.ro.*

27: em.sts.*

28: em.un.*

29: em.vi.*

30: em.htg.hum.*

31: em.htg.inv.*

32: em.htg.other.*

33: em.htg.mus.*

34: em.htg.pln.*

35: em.htg.rod.*

36: em.htg.mam.*

37: em.htg.vrt.*

38: em.sy.*

39: em.htgo.hum.*

40: em.htgo.mus.*

41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	658	55.4	1186	9	HUMTS	D00517 Homo sapien
2	658	55.4	18596	6	AR144965	AR144965 Sequence
3	658	55.4	18596	6	AX050451	AX050451 Sequence
4	658	55.4	18596	6	AX330682	AX330682 Sequence
5	658	55.4	18596	6	AX330906	AX330906 Sequence
6	658	55.4	18596	6	AX335755	AX335755 Sequence
7	658	55.4	18596	6	AX397714	AX397714 Sequence
8	658	55.4	18596	6	AX408943	AX408943 Sequence
9	658	55.4	18596	9	HUMTS1	D00596 Homo sapien
10	545	45.9	152711	9	AP001178	AP001178 Homo sapi
C 11	332	28.0	165742	2	AC020697	AC020697 Homo sapi
12	297	25.0	1536	6	AR103889	AR103889 Sequence
13	297	25.0	1536	6	AX397715	AX397715 Sequence
14	297	25.0	1536	6	AX482020	AX482020 Sequence
15	297	25.0	1536	9	HSTSYN1	X02308 Human mRNA
16	274	23.1	1567	9	AB062290	AB062290 Homo sapi
17	274	23.1	1569	9	BC013919	BC013919 Homo sapi
18	273	23.0	1533	9	BC002567	BC002567 Homo sapi
19	262	22.1	3298	6	AX281791	AX281791 Sequence
20	205	17.3	693	9	AB077207	AB077207 Homo sapi
21	205	17.3	840	9	AB077208	AB077208 Homo sapi
22	205	17.3	942	6	AX069365	AX069365 Sequence
23	205	17.3	942	6	AX397717	AX397717 Sequence
24	170	14.3	249	9	AF279905	AF279905 Homo sapi
C 25	169	14.2	1133	6	AX397748	AX397748 Sequence
C 26	165	13.9	860	9	BC028301	BC028301 Homo sapi
C 27	165	13.9	1131	6	AX397698	AX397698 Sequence
C 28	165	13.9	1161	6	AX397648	AX397648 Sequence
29	142	12.0	165	9	AF279907	AF279907 Homo sapi
30	142	12.0	193	9	AF279906	AF279906 Homo sapi
31	110	9.3	168	9	AF127520	AF127520 Homo sapi
32	108	9.1	535	6	AX341901	AX341901 Sequence
33	82	6.9	140	9	AF127519	AF127519 Homo sapi
C 34	69	5.8	142829	30	AC027512	AC027512 Homo sapi
C 35	69	5.8	159349	9	AC097633	AC097633 Homo sapi
C 36	69	5.8	160693	9	AC124043	AC124043 Homo sapi
C 37	68	5.7	39958	9	AC011542	AC011542 Homo sapi
C 38	68	5.7	104600	9	AP001166	AP001166 Homo sapi
39	68	5.7	107480	9	AL596094	AL596094 Human DNA
40	68	5.7	136283	9	AP005202	AP005202 Homo sapi
C 41	68	5.7	208145	2	AC053504	AC053504 Homo sapi
42	67	5.6	335	9	AF134215	AF134215 Homo sapi
43	64	5.4	176521	9	AC092724	AC092724 Homo sapi
C 44	63	5.3	1938	9	AB060870	AB060870 Macaca fa
C 45	63	5.3	119330	9	AL590076	AL590076 Human DNA

ALIGNMENTS

RESULT 1
LOCUS HUMTS 1186 bp DNA linear PRI 06-NOV-2001
DEFINITION Homo sapiens gene for thymidylate synthase, exon 1, partial cds.
ACCESSION D00517
VERSION D00517.1 GI:220133
KEYWORDS
SOURCE Homo sapiens DNA, clone:lambda HTS-1.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1186)
AUTHORS Takeishi,K., Kaneda,S., Ayusawa,D., Shimizu,K., Gotch,O. and Seno,T.
TITLE Human thymidylate synthase gene: isolation of phage clones which

cover a functionally active gene and structural analysis of the region upstream from the translation initiation codon
T. Biochem. 106 (4), 575-583 (1989)

```
source 1. .18596
/organism="unknown"
BASE COUNT 4521 a 3991 c 4479 g 5605 t
ORIGIN
Query Match 55.4%; Score 658; DB 6; Length 18596;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1178; Conservative 0; Mismatches 8; Indels 1; Gaps 1;
1 GATCGCGCACTGCACTCAGCCTGGGTGAGAGCGAGACTCTCTCTCAAAAAA 60
Db 78 GATCGCGCACTGCACTCAGCCTGGGTGAGAGCGAGACTCTCTCTCAAAAAA 137
2Y 61 AAAAAAGACCGCCAGGCGCTCAAAACAAAAAAGCCTCGAAAAAGCCCTGGGGG 120
Db 138 AAAAAAGACCGCCAGGCGCTCAAAACAAAAAAGCCTCGAAAAAGCCCTGGGGG 197
121 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 180
2Y 198 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 257
Db 181 CGGATCTTGGCTCAGTCAAACTCGCTCCAGGCTGATCAATTTCTTGGCTCAGCC 240
2Y 258 CGGATCTTGGCTCAGTCAAACTCGCTCCAGGCTGATCAATTTCTTGGCTCAGCC 317
Db 241 TCCCAAGTAGCCACACCGCCAGCTAAATTTTGTATTTTGTAGAGACGGGGTTTCAC 300
318 TCCCAAGTAGCCACACCGCCAGCTAAATTTTGTATTTTGTAGAGACGGGGTTTCAC 377
2Y 301 CATGTTCTCAGGCTGGCTGAACTCTCGACCTCAGGTGATCCACCGCCCTCGGCCCCC 360
Db 378 CATGTTCTCAGGCTGGCTGAACTCTCGACCTCAGGTGATCCACCGCCCTCGGCCCCC 437
2Y 361 CAAAGTACTAGGATTAAGGCGTGAGCCACCGCGTCAGCGCCCTCGGGGTTTAAATCA 420
Db 438 CAAAGTACTAGGATTAAGGCGTGAGCCACCGCGTCAGCGCCCTCGGGGTTTAAATCA 497
421 AGTAGAAGAGCTGCAATATACATCTGCTTCCTTCCTTCCTTCCTTCCTTCCTTCCT 480
498 AGTAGAAGAGCTGCAATATACATCTGCTTCCTTCCTTCCTTCCTTCCTTCCTTCCT 556
481 AAATGCAATCNCCTTATAGTTAGTGAAGAACAGATCTCAAAACAGAGTTTGTACAAAG 540
557 AAATGCAATCNCCTTATAGTTAGTGAAGAACAGATCTCAAAACAGAGTTTGTACAAAG 616
541 ACCGACGAAAAGCTGGGAACCTGCTCTGGCTTTAGAGAAAGCGCGGTTCACACAGCGG 600
617 ACCGACGAAAAGCTGGGAACCTGCTCTGGCTTTAGAGAAAGCGCGGTTCACACAGCGG 676
601 TTCCCAAGGGCGCAGTCTCTTCCTCCNGCCACCGCACTGCTTCAGGTTCCCGGGTTCCT 660
677 TTCCCAAGGGCGCAGTCTCTTCCTCCAGCCACCGCACTGCTTCAGGTTCCCGGGTTCCT 736
661 AAGACTCTCAGCTGTGGCCCTTGGGCTCCGTTCTGTGCGCAACCCGTTGCTCTGGTTTC 720
737 AAGACTCTCAGCTGTGGCCCTTGGGCTCCGTTCTGTGCGCAACCCGTTGCTCTGGTTTC 796
721 CCCCTGGCGCAAGCTCTTAGAGCGGGGGCGCGCGAGACCCCGCCAGAGAGAGAGCGG 780
797 CCCCTGGCGCAAGCTCTTAGAGCGGGGGCGCGCGAGACCCCGCCAGAGAGAGAGCGG 856
781 GAGCGCGGAGCGCGCGGGGAAAAGGCGGGGGAGAGGGGTCTTGCACCGCCACTTTGG 840
857 GAGCGCGGAGCGCGCGGGGAAAAGGCGGGGGAGAGGGGTCTTGCACCGCCACTTTGG 916
841 CCTGCTCCGCTCCGCGCGCGCACTTGGCTGCTTCCTTCCTTCCTTCCTTCCTTCCTTC 900
917 CCTGCTCCGCTCCGCGCGCGCACTTGGCTGCTTCCTTCCTTCCTTCCTTCCTTCCTTC 976
901 CCTGCTCCGCTCCGCGCGCGCACTTGGCTGCTTCCTTCCTTCCTTCCTTCCTTCCTTC 960
977 CCTGCTCCGCTCCGCGCGCGCACTTGGCTGCTTCCTTCCTTCCTTCCTTCCTTCCTTC 1036
```

```
QY 961 TTGCCCCCGCCGCGCAAGAGCGGGAGCGCCAGCGCGCTCCGCGCGACGCGGGAGCTGCAG 1020
Db 1037 TTGCCCCCGCCGCGCAAGAGCGGGAGCGCCAGCGCGCTCCGCGCGACGCGGGAGCTGCAG 1096
QY 1021 TACTGGGCGAGATCCAAACATCTCCGCTGGCGGTGAGGAAGGACGACCGCAGGGG 1080
Db 1097 TACTGGGCGAGATCCAAACATCTCCGCTGGCGGTGAGGAAGGACGACCGCAGGGG 1156
QY 1081 ACGGGACACCTGTCGCTATTCGGCATGCGAGCGCGCTACAGCTGAGAGGTGAGCGCGG 1140
Db 1157 ACGGGACACCTGTCGCTATTCGGCATGCGAGCGCGCTACAGCTGAGAGGTGAGCGCGG 1216
QY 1141 GGCCCTTGGCGGACGCGGTGGCGGGAAGGAGGAGCGCGCGCTGGGGA 1187
Db 1217 GGCCCTTGGCGGACGCGGTGGCGGGAAGGAGGAGCGCGCGCTGGGGA 1263
RESULT 3
AX050451 18596 bp DNA linear PAT 12-JAN-2001
LOCUS Sequence 11 from Patent WO0071754.
DEFINITION AX050451
ACCESSION AX050451
VERSION AX050451.1 GI:12226657
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 18596)
AUTHORS Johnson, W.G. and Stenroos, E.S.
TITLE Methods for diagnosing, preventing, and treating developmental
disorders due to a combination of genetic and environmental factors
JOURNAL Patent: WO 0071754-A 11 30-NOV-2000;
University of Medicine and Dentistry of New Jersey (US)
FEATURES
source
1. .18596
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 4521 a 3991 c 4479 g 5605 t
ORIGIN
Query Match 55.4%; Score 658; DB 6; Length 18596;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1178; Conservative 0; Mismatches 8; Indels 1; Gaps 1;
QY 1 GATCGCGCACTGCACTCAGCCTGGGTGAGAGCGAGACTCTCTCTCAAAAAA 60
Db 78 GATCGCGCACTGCACTCAGCCTGGGTGAGAGCGAGACTCTCTCTCAAAAAA 137
QY 61 AAAAAAGACCGCCAGGCGCTCAAAACAAAAAAGCCTCGAAAAAGCCCTGGGGG 120
Db 138 AAAAAAGACCGCCAGGCGCTCAAAACAAAAAAGCCTCGAAAAAGCCCTGGGGG 197
QY 121 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 180
Db 198 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 257
QY 181 CGGATCTTGGCTCAGTCAAACTCGCTCCAGGCTGATCAATTTCTTGGCTCAGCC 240
Db 258 CGGATCTTGGCTCAGTCAAACTCGCTCCAGGCTGATCAATTTCTTGGCTCAGCC 317
QY 241 TCCCAAGTAGCCACACCGCCAGCTAAATTTTGTATTTTGTAGAGACGGGGTTTCAC 300
Db 318 TCCCAAGTAGCCACACCGCCAGCTAAATTTTGTATTTTGTAGAGACGGGGTTTCAC 377
QY 301 CATGTTCTCAGGCTGGCTGAACTCTCGACCTCAGGTGATCCACCGCCCTCGGCCCCC 360
Db 378 CATGTTCTCAGGCTGGCTGAACTCTCGACCTCAGGTGATCCACCGCCCTCGGCCCCC 437
QY 361 CAAAGTACTAGGATTAAGGCGTGAGCCACCGCGTCAGCGCCCTCGGGGTTTAAATCA 420
Db 438 CAAAGTACTAGGATTAAGGCGTGAGCCACCGCGTCAGCGCCCTCGGGGTTTAAATCA 497
```


961 TTGCCCCCGCCGACAGAGCGGAGCGCGGCTCCGCGCCGACGGGAGCTGAG 1020
1037 TTGCCCCCGCCGACAGAGCGGAGCGCGGCTCCGCGCCGACGGGAGCTGAG 1096
1021 TACCTGGGGCAGATCAACACATCTCCGCTGGCGGCTCAGGAAGGACGACCGACGGC 1080
1097 TACTGGGGCAGATCAACACATCTCCGCTGGCGGCTCAGGAAGGACGACCGACGGC 1156
1081 ACCGGACCTCTGCGTATTTGGCATGATGAGCGCGCTACAGCTTGAAGGTTGACGCGCG 1140
1157 ACCGGACCTCTGCGTATTTGGCATGATGAGCGCGCTACAGCTTGAAGGTTGACGCGCG 1216
1141 GGCCCCGCGGAGCGGCTGGCGGAGAGGAGCGCGCGCTGGGGA 1187
1217 GGCCCCGCGGAGCGGCTGGCGGAGAGGAGCGCGCGCTGGGGA 1263

RESULT 5
AX330906
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES

AX330906 18596 bp DNA linear PAT 09-JAN-2002
Sequence 1415 from Patent WO0194629.
AX330906
AX330906.1 GI:18103885
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
Horrihan, S., Soppet, D.R. and Weaver, Z.
Cancer gene determination and therapeutic screening using signature
gene sets
Patent: WO 0194629-A 1415 13-DEC-2001;
Avalon Pharmaceuticals (US)
Location/Qualifiers

1. 18596
/organism="Homo sapiens"
/db_xref="taxon:9606"
4521 a 3991 c 4479 g 5605 t

BASE COUNT 4521 a 3991 c 4479 g 5605 t
ORIGIN

Query Match 55.4%; Score 658; DB 6; Length 18596;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1178; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

1 GATCGCCCACTGCACTCCAGCGCTGGGAGAGAGGAGAGCTCTGCTCAAAAAA 60
78 GATCGCCCACTGCACTCCAGCGCTGGGAGAGAGGAGAGCTCTGCTCAAAAAA 137
61 AAAAAGACCGCCAGGCTCAACAAAAAACCCTCGAAAAAGCCCTGGCGGCTCTTTT 120
138 AAAAAGACCGCCAGGCTCAACAAAAAACCCTCGAAAAAGCCCTGGCGGCTCTTTT 197
121 TTTTCTTTTCTTTTCTTTTGGGACAGTCTGCTCTGCGCCAGGCTGAGTACAAATGGT 180
198 TTTTCTTTTCTTTTCTTTTGGGACAGTCTGCTCTGCGCCAGGCTGAGTACAAATGGT 257
181 CGGATCTTGCTCACTGCAACCTCTGCTCCAGGTTCAAGCAATTTCTGCTCAGCC 240
258 CGGATCTTGCTCACTGCAACCTCTGCTCCAGGTTCAAGCAATTTCTGCTCAGCC 317
241 TCCCAAGTAGCAGCAGCCAGCTCAATTTTGTANTTTTGTAGAGAGCGGGGTTTCA 300
318 TCCCAAGTAGCAGCAGCCAGCTCAATTTTGTANTTTTGTAGAGAGCGGGGTTTCA 377
301 CATGTTCTCAGGCTGCTGCACTCTGCTCAGTCACTGATCCAGCCGCTCGGCCCCC 360
378 CATGTTCTCAGGCTGCTGCACTCTGCTCAGTCACTGATCCAGCCGCTCGGCCCCC 437
361 CAAAGTACTAGGATTAACGGCTGAGCAGCAGCGCTGAGCGGCTGGGTTTAAATCA 420
438 CAAAGTACTAGGATTAACGGCTGAGCAGCAGCGCTGAGCGGCTGGGTTTAAATCA 497

QY 421 AGTAGAAAAAGCTGCATTATACCACTTCTTCTGNTTTCAGTGAGAACGAAATGG 480
Db 498 AGTAGAAAAAGCTGCATTATACCACTTCTTCTGNTTTCAGTGAGAACGAAATGG 556
QY 481 AAATGCAAAATCCTTATTTAGTTTGTAGAAACAGATCTCAAAACAGCAGTTTGTGACAAG 540
Db 557 AAATGCAAAATCCTTATTTAGTTTGTAGAAACAGATCTCAAAACAGCAGTTTGTGACAAG 616
QY 541 ACCCAGAAAACGTTGGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
Db 617 ACCCAGAAAACGTTGGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 676
QY 601 TTCCCAAAAGGCGCAGTCTCTCCNGCCACCGCACCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Db 677 TTCCCAAAAGGCGCAGTCTCTCTCCAGCCACCGCACCTGCTGCTGCTGCTGCTGCTGCT 736
QY 661 AAGACTCTCAGCTGTGGCCCTGGGCTCCGTTCTGTGCCACACCCCTGCTGCTGCTGCTGCT 720
Db 737 AAGACTCTCAGCTGTGGCCCTGGGCTCCGTTCTGTGCCACACCCCTGCTGCTGCTGCTGCT 796
QY 721 CCCCTGGCGCAGCTCTCTAGAGCGGGGCGCGCGACCCCGCTGAGAGAGAGAGCGG 780
Db 797 CCCCTGGCGCAGCTCTCTAGAGCGGGGCGCGCGACCCCGCTGAGAGAGAGAGCGG 856
QY 781 GAGCGCGGAGCGCGCGGGAAGGCGCGGGAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Db 857 GAGCGCGGAGCGCGCGGGAAGGCGCGGGAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 916
QY 841 CTTGCTCTCCGCTCCGCGCGCGCACCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
Db 917 CTTGCTCTCCGCTCCGCGCGCGCACCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 976
QY 901 CTTGCTCTCCGCTCCGCGCGCGCACCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
Db 977 CTTGCTCTCCGCTCCGCGCGCGCACCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1036
QY 961 TTGCCCCCGCCGACAGAGCGGAGCGCGGAGCGCGCTCCGCGCCGACCGGAGCTGCGAG 1020
Db 1037 TTGCCCCCGCCGACAGAGCGGAGCGCGGAGCGCGCTCCGCGCCGACCGGAGCTGCGAG 1096
QY 1021 TACCTGGGGCAGATCAACACATCTCCGCTGGCGGCTCAGGAAGGACGACCGACGGGCG 1080
Db 1097 TACTGGGGCAGATCAACACATCTCCGCTGGCGGCTCAGGAAGGACGACCGACGGGCG 1156
QY 1081 ACCGGACCTCTGCGTATTTGGCATGATGAGCGCGCTACAGCTTGAAGGTTGACGCGCG 1140
Db 1157 ACCGGACCTCTGCGTATTTGGCATGATGAGCGCGCTACAGCTTGAAGGTTGACGCGCG 1216
QY 1141 GGCCCCGCGGAGCGGCTGGCGGAGAGGAGCGCGCGCTGGGGA 1187
Db 1217 GGCCCCGCGGAGCGGCTGGCGGAGAGGAGCGCGCGCTGGGGA 1263

RESULT 6
AX335755
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES

AX335755 18596 bp DNA linear PAT 09-JAN-2002
Sequence 6264 from Patent WO0194629.
AX335755
AX335755.1 GI:18126474
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
Horrihan, S., Soppet, D.R. and Weaver, Z.
Cancer gene determination and therapeutic screening using signature
gene sets
Patent: WO 0194629-A 6264 13-DEC-2001;
Avalon Pharmaceuticals (US)
Location/Qualifiers

/organism="Homo sapiens"	
/db xref="taxon:9606"	
/note="EMBL/GenBank Accession No. D00596"	
BASE COUNT	4521 a 3991 c 4479 g 5605 t
ORIGIN	
Query Match	55.4%; Score 658; DB 6; Length 18596;
Best Local Similarity	99.2%; Pred. No. 0;
Matches 1178; Conservative	0; Mismatches 8; Indels 1; Gaps
Qy	1 GATCGCGCACTGCACTCCAGCCTGGGTGAGAGCGAGACTCTCTCTCAAAAAA 60
Db	78 GATCGCGCACTGCACTCCAGCCTGGGTGAGAGCGAGACTCTCTCTCAAAAAA 137
Qy	61 AAAAAGACCGCAGCGCTCAAAACAAAACCTCGSAAAAGCCCTCGCGGTCTTTT 120
Db	138 AAAAAGACCGCAGCGCTCAAAACAAAACCTCGSAAAAGCCCTCGCGGTCTTTT 197
Qy	121 TTTTCTTTTCTTTTCTTTTGGGACGCTCTTGCTCTGTGCCCCAGCTGGAGTCAATGGT 180
Db	198 TTTTCTTTTCTTTTCTTTTGGGACGCTCTTGCTCTGTGCCCCAGCTGGAGTCAATGGT 257
Qy	181 CGGATCTTGCTCACTGCAACACTCTGCCCTCCAGGTTCAAGCAATTCTTGCTTCAGCC 240
Db	258 CGGATCTTGCTCACTGCAACACTCTGCCCTCCAGGTTCAAGCAATTCTTGCTTCAGCC 317
Qy	241 TCCCAAGTAGCACACACGCCCCACTTAATTTTGTATTTTACTAGACAGCGGGTTTCAC 300
Db	318 TCCCAAGTAGCACACACGCCCCACTTAATTTTGTACTTTTAGTAGACCGGGTTTCAC 377
Qy	301 CATGTTCTCCAGCGTGTCTGAACTCTGTGACCTCAGGTGATCCACCGCTCGGCCGCC 360
Db	378 CATGTTCTCCAGCGTGTCTGAACTCTGTGACCTCAGGTGATCCACCGCTCGGCCGCC 437
Qy	361 CAAAGTACTAGGATTTACAGCGGTGAGCCACCGCGCTCCAGCGCTCTGGCGGTTTTAAATCA 420
Db	438 CAAAGTACTAGGATTTACAGCGGTGAGCCACCGCGCTCCAGCGCTCTGGCGGTTTTAAATCA 497
Qy	421 AGTAGAAAAGCTGCATATACCACTTGCTTCNGTTGNTTTCAGTCAGAACCAAGAAATGG 480
Db	498 AGTAGAAAAGCTGCATATACCACTTGCTTCGGTTGC-TTCAGTCAGAACCAAGAAATGG 556
Qy	481 AAATGCAAAATCMTTATTAGTTGTAGGAAAACAGATCTCAAAACAGCAGTTTGTGNGACAAG 540
Db	557 AAATGCAAAATCCTTTATTAGTTGTAGGAAAACAGATCTCAAAACAGCAGTTTGTGTCAGAG 616
Qy	541 ACCCGAGGAAAACGTGGGAACCTGTGCTGCTGGCTTAGAGGAAGCCGCGTCGACAGAGG 600
Db	617 ACCCGAGGAAAACGTGGGAACCTGTGCTGCTGGCTTAGAGGAAGCCGCGTCGACAGAGG 676
Qy	601 TTCTCCAAAGGGCGAGTCTTTCCTCCAGCACCGACCTGNTTCCAGTTCCCGGTTTCTCT 660
Db	677 TTCTCCAAAGGGCGAGTCTTTCCTCCAGCACCGACCTGNTTCCAGTTCCCGGTTTCTCT 736
Qy	661 AAGACTCTCAGCTGTGGCCCTCGGGCTTCGGTCTGTGTCACACCCCGTGGCTCTCTGCGTTTC 720
Db	737 AAGACTCTCAGCTGTGGCCCTCGGGCTTCGGTCTGTGTCACACCCCGTGGCTCTCTGCGTTTC 796
Qy	721 CCCCCTGGCGACGCTCTTAGAGCGGGGGCCCGCGACCCCGCGAGCAGGAAGAGCGC 780
Db	797 CCCCCTGGCGACGCTCTTAGAGCGGGGGCCCGCGACCCCGCGAGCAGGAAGAGCGC 856
Qy	781 GAGCGCGGAGCGCGCGGAAAAGCGCGCGGAAGGGGTCCTGCCACCGCGCCACTTGG 840
Db	857 GAGCGCGGGAACGCGCGGGAAGGGGCTCTGCCACCGCGCCACTTGG 916
Qy	841 CTTGCTCTGTCGCGCGCGCACTTGCGCTGCTCTGCTCCCGCGCGGCCACTTTCGCTG 900
Db	917 CTTGCTCTGTCGCGCGCGCACTTGCGCTGCTCTGCTCCCGCGCGGCCACTTTCGCTG 976
Qy	901 CTTGCTCTCCCGCGCGCGCGCACTTGCTTGCGCGGGCTCGGAGCTGCGCGCGCGGCC 960
Db	977 CTTGCTCTCCCGCGCGCGCGCACTTGCTTGCGCGGGCTCGGAGCTGCGCGCGCGGCC 1036

438	CAAAGTACTAGGATTACAGGCGTAGGACACCGCGTCCAGCGCCCTCGGCGTTTATCA	497
441	AGTACAAAAGCTGATATTAACACTTTCCTTCNGTTGCTNTTCAGTGCAGAACGAGAATGG	480
444	AGTAGAAAAGCTGCATTATACCACTTCCTTCGGTTGC-TTCAGTGGAAACAGAGATGG	556
447	AAATGCAAAATCNCCTTTATAGTTGTAGGAAACAGATCTCAAACAGCAGTTTGTGNGACAAG	540
450	AAATGCAAAATCCCTTATTAGTTGTAGGAAACAGATCTCAAACAGCAGTTTGTGACAAG	616
453	ACCGACGAAAACGTGGGAACGTGCTCTGCTGCTTGTAGAGAGGCGCGTCCAGACACGG	600
456	ACCGACGAAAACGTGGGAACGTGCTCTGCTGCTTGTAGAGAGGCGCGTCCAGACACGG	676
459	TTCCCAAGGGCGCACTCTCCNGCCACGCACTGCTCCAGGTTCCCGGGTTTCCT	660
462	TTCCCAAGGGCGCACTCTCCAGCCACCGACCTGATCAGGTTCCCGGGTTTCCT	736
465	AAGACTCTCAGCTGTGGCCCTCGGCTCTGTGCCACACCCCTGGCTCTCGCGTTTC	720
468	AAGACTCTCAGCTGTGGCCCTCGGCTCTGTGCCACACCCCTGGCTCTCGCGTTTC	796
471	CCCTGGCGCA CGCTCTCTAGAGCGGGGCGCGCGACCCCGCGACGAGGAGGCG	780
474	CCCTGGCGCA CGCTCTCTAGAGCGGGGCGCGCGACCCCGCGACGAGGAGGCG	856
477	GAGCGCGGAGCGCCCGCGGAAAAGGCGCGGAGAGGGTCTTGCACCGCGCCACTTGG	840
480	GAGCGCGGAGCGCCCGCGGAAAAGGCGCGGAGAGGGTCTTGCACCGCGCCACTTGG	916
483	CCTGCCTCCGTCGCGCGCGGCACTTGGCTGTGCTCCGTCGCGCGCGCCACTTGCCTG	900
486	CCTGCCTCCGTCGCGCGCGGCACTTGGCTGTGCTCCGTCGCGCGCGCCACTTGCCTG	976
489	CCTGCCTCCGTCGCGCGCGGCACTTGGCTGTGCTCCGTCGCGCGCGCCACTTGCCTG	960
492	CCTGCCTCCGTCGCGCGCGGCACTTGGCTGTGCTCCGTCGCGCGCGCCACTTGCCTG	1036
495	TTGCCCCCGCGCGCACAGGAGCGGAGCCCGACGCGCTCCGCGCACGCGGAGCTGCAG	1020
498	TTGCCCCCGCGCGCACAGGAGCGGAGCCCGACGCGCTCCGCGCACGCGGAGCTGCAG	1096
501	TACTGGGGCAGATCCATCATCTCCGTCGCGGCTCAGAGAGGAGCAGCGCAGCGGC	1080
504	TACTGGGGCAGATCCATCATCTCCGTCGCGGCTCAGAGAGGAGCAGCGCAGCGGC	1156
507	ACGGCACCTGTGCGGTATTCGGCATACAGGCGCGCTACAGCCTGAGAGTGAAGCGCG	1140
510	ACGGCACCTGTGCGGTATTCGGCATACAGGCGCGCTACAGCCTGAGAGTGAAGCGCG	1216
513	GGCCCTCGCGGACCGGTGCGGGAAGGAGGCGCGGCTGGGGA	1187
516	GGCCCTCGCGGACCGGTGCGGGAAGGAGGCGCGGCTGGGGA	1263
RESULT 8		
LOCUS	AX408943	
DEFINITION	Sequence 1590 from Patent WO229103.	
ACCESSION	AX408943	
VERSION	AX408943.1	GI:21441648
KEYWORDS	human.	
SOURCE	Homo sapiens	
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
AUTHORS	Alvares, C., Horne, D., Peres-da-Silva, S. and Vockley, J. G.	
TITLE	Gene expression profiles in liver cancer	
JOURNAL	Patent: WO 0229103-A 1590 11-APR-2002;	
FEATURES	GENE LOGIC INC (US)	
source	Location/Qualifiers	
	1..18596	

61 AAAAGACCGCCAGGGCTCAAAACAAAACCTCGGAAAGCCCTGCGGCTCTTTTTTTTT 120
 138 AAAAGACCGCCAGGGCTCAAAACAAAACCTCGGAAAGCCCTGCGGCTCTTTTTTTTT 197
 121 TTTTCTTTTCTTTTCTTTTGGGACAGTCTTGCTCTGTGCGCCAGGCTGGAGTACAATGGT 180
 198 TTTTCTTTTCTTTTCTTTTGGGACAGTCTTGCTCTGTGCGCCAGGCTGGAGTACAATGGT 257
 181 CGGATCTTGGCTCACTGCAACCTCTGCTCTCCAGGTTCAAGCAATTTCTTCTGCTCAGCC 240
 258 CGGATCTTGGCTCACTGCAACCTCTGCTCTCCAGGTTCAAGCAATTTCTTCTGCTCAGCC 317
 241 TCCCAAGTAGCCACCGCCAGCTAAATTTTGTANTTTAGTAGAGAGCGGGGTTTCCAC 300
 318 TCCCAAGTAGCCACCGCCAGCTAAATTTTGTACTTTTAGTAGAGAGCGGGGTTTCCAC 377
 301 CATGTTGTCCAGGCTGTCTNGAATCTCTGACCTCAGTGATCCACCGCTCGACCC 360
 378 CATGTTGTCCAGGCTGTCTTGAATCTCTGACCTCAGTGATCCACCGCTCGACCC 437
 361 CAAAGTACTAGGATTAAGGCTGAGCCAGCGCTCCAGCGCTCGCGGTTTAAATCA 420
 438 CAAAGTACTAGGATTAAGGCTGAGCCAGCGCTCCAGCGCTCGCGGTTTAAATCA 497
 421 AGTAGAAGCTGATTAATACCACTTGCTTCTGCTGCTTGTAGAGAGCGGCTCGACAGAG 480
 498 AGTAGAAGCTGATTAATACCACTTGCTTCTGCTGCTTGTAGAGAGCGGCTCGACAGAG 556
 481 AAATGCAATCTCTTATAGTTGTAGGAAACAGATCTCAACACGAGTTTGTGACAAAG 540
 557 AAATGCAATCTCTTATAGTTGTAGGAAACAGATCTCAACACGAGTTTGTGACAAAG 616
 541 ACCGAGGAAACGCTGGAACCTGTGCTGTGCTGTAGAGAGCGGCTCGACAGAGCG 600
 617 ACCGAGGAAACGCTGGAACCTGTGCTGTGCTGTAGAGAGCGGCTCGACAGAGCG 676
 601 TTCCAAAGGCGCAGTCTTCCGACCGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
 677 TTCCAAAGGCGCAGTCTTCCGACCGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 736
 661 AAGACTCTCAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 737 AAGACTCTCAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 796
 721 CCCCTGCGCAGCTCTCTAGAGCGGGGCGCGCGACCTGCTGCTGCTGCTGCTGCTGCTGCT 780
 797 CCCCTGCGCAGCTCTCTAGAGCGGGGCGCGCGACCTGCTGCTGCTGCTGCTGCTGCTGCT 856
 781 GAGCGCGGAGCGCGCGGAAAGCGCGGGAAGGGGCTCTGCTGCTGCTGCTGCTGCTGCTGCT 840
 857 GAGCGCGGAGCGCGCGGAAAGCGCGGGAAGGGGCTCTGCTGCTGCTGCTGCTGCTGCTGCT 916
 841 CTTGCTCTCCGTCGCGCGGCACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
 917 CTTGCTCTCCGTCGCGCGGCACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 976
 901 CTTGCTCTCCGTCGCGCGGCACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
 977 CTTGCTCTCCGTCGCGCGGCACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1036
 961 TTGCCCCCGCGCGCAGAGCGGGAAGCGCGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
 1037 TTGCCCCCGCGCGCAGAGCGGGAAGCGCGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1096
 1021 TACTTGGGCGAGATCCAAACATCTTCCCTCTCGCGGCTCAGGAAAGGACGACGCGCG 1080
 1097 TACTTGGGCGAGATCCAAACATCTTCCCTCTCGCGGCTCAGGAAAGGACGACGCGCG 1156
 1081 ACCGCAACCTTGTGCTGTAATTCGCAATGACGCGCGCTTACGCTTGAAGGTGACGCGCG 1140
 1157 ACCGCAACCTTGTGCTGTAATTCGCAATGACGCGCGCTTACGCTTGAAGGTGACGCGCG 1216
 1141 GGCCTCTCGGAGCGGCTGCGGGAAGGAGGAGCGCGCTGCGGA 1187

[illegible]

DEFINITION Sequence 11 from Patent WO02057489.
ACCESSION AX482020
VERSION AX482020.1 GI:22316743
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Danenberg, K.D.
TITLE Method of determining a chemotherapeutic regimen based on ercc1 and ts expression
JOURNAL Patent: WO 02057489-A 11 25-JUL-2002;
RESPONSE GENETICS INC (US)
FEATURES Location/Qualifiers
source 1..1536
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 390 a 369 c 399 g 378 t
ORIGIN
Query Match 25.0%; Score 297; DB 6; Length 1536;
Best Local Similarity 100.0%; Pred. No. 4.8e-157;
Matches 297; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 833 CCACCTGGCTGCTCCGTCGCGCGCGCCACCTTGGCTGCTCCGTCGCGCGCCAC 892
Db 14 CCACCTGGCTGCTCCGTCGCGCGCGCCACCTTGGCTGCTCCGTCGCGCGCCAC 73
QY 893 TTGCGCTGCTCCGTCGCGCGCGCCACCTTGGCTGCTCCGTCGCGCGCCAC 952
Db 74 TTGCGCTGCTCCGTCGCGCGCGCCACCTTGGCTGCTCCGTCGCGCGCCAC 133
QY 953 GCGGCGCTTGGCG 1012
Db 134 GCGGCGCTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 193
QY 1013 AGCTGCAGTACCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1072
Db 194 AGCTGCAGTACCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 253
QY 1073 GCACGGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGAC 1129
Db 254 GCACGGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGAC 310
RESULT 15
HSTSVM1 1536 bp mRNA linear PRI 12-SEP-1993
LOCUS Human mRNA for thymidylate synthase (EC 2.1.1.45).
DEFINITION Human mRNA for thymidylate synthase (EC 2.1.1.45).
ACCESSION X02308
VERSION X02308.1 GI:37478
KEYWORDS inverted repeat; synthetase; tandem repeat.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1536)
AUTHORS Takeishi, K., Kaneda, S., Ayusawa, D., Shimizu, K., Gotoh, O. and Seno, T.
TITLE Nucleotide sequence of a functional cDNA for human thymidylate synthase
JOURNAL Nucleic Acids Res. 13 (6), 2035-2043 (1985)
MEDLINE 85215597
PUBMED 2987839
COMMENT Data kindly reviewed (22-OCT-1985) by Seno T.
FEATURES Location/Qualifiers
source 1..1536
/organism="Homo sapiens"
/db_xref="taxon:9606"
misc_feature 14..103
/notes="triple tandemly repeated elements"
repeat_region 14..16

misc_feature 35..103
/note="direct repeat 1"
misc_feature 35..69
/note="pot. stem-loop structure"
repeat_unit 35..41
/note="pot. stem-loop structure"
repeat_region 42..44
/note="inverted repeat A"
misc_feature 63..103
/note="direct repeat 1"
repeat_unit 63..69
/note="inverted repeat A"
repeat_region 70..72
/note="direct repeat 1"
repeat_unit 97..103
/note="inverted repeat A"
repeat_region 104..106
/note="direct repeat 1"
CDS 106..1047
/note="Chymidylate synthase (aa 1-313)"
/codon_start=1
/protein_id="CAA26178.1"
/db_xref="GI:37479"
/db_xref="SWISS-PROT:P04818"
/translation="MPVAGSELPRRLPPAQRDAPRRPHGLQVLQLOIHLRCG
VKDRDTGTLGVFGNARYSLRDEFFLLTKRVFKVGLLELLMFTIKGSTNAKELS
SKGVKIDWANGSRDLDLGFSTREEDLGPVYGFQWRFHFAEYRWESDYSGQVDO
LQVDTITKIPDDRIIMCANWPRDLPLMALPPCHALCOFYVNVSELSOLYQSSGD
MGLGVFNFIASVALLTYIAHITGLKPGDFIHTGLDADHIYLNIEPLKIQLOREPRPF
PKRLIRKVEKIDFKAEDFQIEGYNPHPTIKMEAV"
misc_feature 1231..1236
/note="pot. polyadenylation signal"
misc_feature 1519..1524
/note="pot. polyadenylation signal"
polyA_site 1536
/note="polyadenylation site"
BASE COUNT 390 a 369 c 399 g 378 t
ORIGIN
Query Match 25.0%; Score 297; DB 9; Length 1536;
Best Local Similarity 100.0%; Pred. No. 4.8e-157;
Matches 297; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 833 CCACCTGGCTGCTCCGTCGCGCGCGCCACCTTGGCTGCTCCGTCGCGCGCCAC 892
Db 14 CCACCTGGCTGCTCCGTCGCGCGCGCCACCTTGGCTGCTCCGTCGCGCGCCAC 73
QY 893 TTGCGCTGCTCCGTCGCGCGCGCCACCTTGGCTGCTCCGTCGCGCGCCAC 952
Db 74 TTGCGCTGCTCCGTCGCGCGCGCCACCTTGGCTGCTCCGTCGCGCGCCAC 133
QY 953 GCGGCGCTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1012
Db 134 GCGGCGCTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 193
QY 1013 AGCTGCAGTACCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1072
Db 194 AGCTGCAGTACCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 253
QY 1073 GCACGGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGAC 1129
Db 254 GCACGGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGAC 310

Search completed: November 29, 2002, 02:33:58
Job time : 3313.34 secs


```
b 361 ATCTAAAGAGCTGTCTTCCAGGAGTGAATCTGGGATGCCAATGGAATCCCGAGACT 420
y 455 TTTTGGACAGCTGGGATCTTCCACAGAGAGGAGGACTTGGGCCCGAGTTATGGCT 514
b 421 TTTTGGACAGCTGGGATCTTCCACAGAGAGGAGGACTTGGGCCCGAGTTATGGCT 480
y 515 TCCAGTGGAGGCAATTTTGGGCGAGATPACAGAGATATGGAATCAGATTATTCAGGACAGS 574
b 481 TCCAGTGGAGGCAATTTTGGGCGAGATPACAGAGATATGGAATCAGATTATTCAGGACAGS 540
y 575 GAGTTGACCACTGCAAGAGTGTATGACACATCAAAACCAACCTTGACGACAGAGAA 634
b 541 GAGTTGACCACTGCAAGAGTGTATGACACATCAAAACCAACCTTGACGACAGAGAA 600
y 635 TCATCATGTGGCTTGAATCCAGAGATCTTCTCTGATGGCGCTGCCCTCCATGCCATG 694
b 601 TCATCATGTGGCTTGAATCCAGAGATCTTCTCTGATGGCGCTGCCCTCCATGCCATG 660
y 695 CCTCTGCCAGTTCCTATGTGGTGAACAGTGTGCTGCTGCCAGCTGTACAGAGATCGG 754
b 661 CCTCTGCCAGTTCCTATGTGGTGAACAGTGTGCTGCTGCCAGCTGTACAGAGATCGG 720
y 755 GAGACATGGGCTCGGTGTGCTTTCAACATGCGCAGCTACGCCCTGCTCACTGATCA 814
b 721 GAGACATGGGCTCGGTGTGCTTTCAACATGCGCAGCTACGCCCTGCTCACTGATCA 780
y 815 TTGCGCATACACGCGGCTGAAAGCTGAAATTCAGCTTCAGCGAGAACCCAGCTTTCC 934
b 841 TTGCGCATACACGCGGCTGAAAGCTGAAATTCAGCTTCAGCGAGAACCCAGCTTTCC 900
y 935 CAAAGCTCAGAGATCTTCCGAAAGTGTGAGAAATGATGACTTCAAAGCTGAAAGCTTTC 994
b 901 CAAAGCTCAGAGATCTTCCGAAAGTGTGAGAAATGATGACTTCAAAGCTGAAAGCTTTC 960
y 995 AGATTGAAGGGTACATCCGATCCCACTTAAATGAAATGGCTTTTGGGTTGCTT 1054
b 961 AGATTGAAGGGTACATCCGATCCCACTTAAATGAAATGGCTTTTGGGTTGCTT 1020
```

```
RESULT 5
L562750/c 1021 bp mRNA linear EST 16-FEB-2001
OCUS AL562750 LTI_NFL003_NBC3 Homo sapiens cDNA clone CS0DC018YA14 3
DEFINITION prime, mRNA sequence.
CCESSTON AL562750
ERSION AL562750.1 GI:12911478
EYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1. (bases 1 to 1021)
Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
1. 1021
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DC018YA14"
/journal="LTI_NFL003_NBC3"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
```

FEATURES
source

/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime and enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 276 a 230 c 307 t 5 others

ORIGIN

Query Match 61.4%; Score 943.8; DB 9; Length 1021;
Best Local Similarity 97.8%; Pred. No. 5.3e-235;
Matches 962; Conservative 3; Mismatches 18; Indels 1; Gaps 1;

```
QY 529 TTTGGGCGAAGATACAGAGATATGGAATCAGATTATTCAGGACAGGAGTTGACCACTG 588
Db 984 TTTGGGCGAAGAAACAGAGATATGAAATCAGATTATTCAGAACAGGAGTTGACCACTG 925
QY 589 CAAAGAGTGTATGACACCATCAAAACCAACCTTGACGACAGAGATCATCATGTGCGCT 648
Db 924 CAAAGAGTGTATGACACCATCAAAACCAACCTTGACGACAGAGATCATCATGTGCGCT 865
QY 649 TGGAAAT-CCAGAGATCTTCTCTGATGGCGCTGCTCCATGCCATGCCCTCTGCCAGTT 707
Db 864 TGGAAATCCAGAGATCTTCTCTGATGGCGCTGCTCCATGCCATGCCCTCTGCCAGTT 805
QY 708 CTATGTGGTGAACAGTGTGCTTCCAGCTGTACAGAGATCGGAGACATGGGCT 767
Db 804 CTATGTGGTGAACAGTGTGCTTCCAGCTGTACAGAGATCGGAGAAACATGGGCT 745
QY 768 CGGTGTGCTTTTCAACATGCGCAGCTACGCGCTGCTCGTACATGATTCGCACTCAC 827
Db 744 CGGTGTGCTTTTCAACATGCGCAGCTACGCGCTGCTCGTACATGATTCGCACTCAC 685
QY 828 GGGCTTGAAGCCAGTGTACTTATACACACTTTGGAGATGCACATATTACCTGAATCA 887
Db 684 GGGCTTGAAGCCAGTGTACTTATACACACTTTGGAGATGCACATATTACCTGAATCA 625
QY 888 CATTGAGCCACTGAAATTCAGCTTTCAGCGAGAACCCAGACCTTTCCAAAGCTCAGGAT 947
Db 624 CATTGAGCCACTGAAATTCAGCTTTCAGCGAGAACCCAGACCTTTCCAAAGCTCAGGAT 565
QY 948 TCTTCCAAAAGTTGAGAAATTCAGCTTCAAAGCTGAAGACTTTCAGATTGAAGGTA 1007
Db 564 TCTTCCAAAAGTTGAGAAATTCAGCTTCAAAGCTGAAGACTTTCAGATTGAAGGTA 505
QY 1008 CAATCCGCACTCCACTATTAAATGGAATGGCTGTTTAGGGTGTCTTTCAAAGGAGCTNG 1067
Db 504 CAATCCGCACTCCACTATTAAATGGAATGGCTGTTTAGGGTGTCTTTCAAAGGAGCTCG 445
QY 1068 AAGGATATGTCAGTCTTTAGGGGTTGGGCTGGATGCGGAGGTAAAAGTTCTTTTGGCTC 1127
Db 444 AAGGATATGTCAGTCTTTAGGGGTTGGGCTGGATGCGGAGGTAAAAGTTCTTTTGGCTC 385
QY 1128 TAAAGAAAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1187
Db 384 TAAAGAAAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 325
QY 1188 GATGTGTCACCTGCAAAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 1247
Db 324 GATGTGTCACCTGCAAAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 265
QY 1248 TAACTCACTGAGGATCTGCAATGCTGAGGTATGACAAAGTGAAGGAGATGAAATG 1307
Db 264 TAACTCACTGAGGATCTGCAATGCTGAGGTATGACAAAGTGAAGGAGATGAAATG 205
QY 1308 TATGTCTCTTAGCAAAAGTGAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 1367
Db 204 TATGTCTCTTAGCAAAAGTGAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 145
```

QY 1368 TGGTGAATTCACAGCTATTTTGGAAATATTTTAGAATATTTTAAAGATTTTCAACGC 1427
 Db 144 TGGTGAATTCACAGCTATTTTGGAAATATTTTAGAATATTTTAAAGATTTTCAACGC 85

QY 1428 TATTCCCTCAATCTGAGGAGCTGAGTACACCATCGATCATCATGATGAGTGTGGTTA 1487
 Db 84 TATTCCCTCAATCTGAGGAGCTGAGTACACCATCGATCATCATGATGAGTGTGGTTA 25

QY 1488 TGAACCTTTAAGTGTGTTTATATG 1511
 Db 24 TGAACCTTTAAGTGTGTTTATG 1

RESULT 6
 AL513730 978 bp mRNA linear EST 13-FEB-2001
 LOCUS AL513730 LTI NFL006 PL2 Homo sapiens cDNA clone CLOB0A082B11 5
 DEFINITION prime, mRNA sequence.
 ACCESSION AL513730
 VERSION AL513730.1 GI:12777224
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 978)
 Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 CONTACT: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
 source
 1..978
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CLOB0A082B11"
 /clone_lib="LTI NFL006 PL2"
 /tissue_type="placenta"
 /note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 238 a 258 c 263 g 211 t 8 others
 ORIGIN

Query Match 60.9%; Score 934.8; DB 9; Length 978;
 Best Local Similarity 98.6%; Pred. No. 1.2e-232;
 Matches 966; Conservative 7; Mismatches 4; Indels 3; Gaps 3;

QY 80 TGCCTCCGTCCTCCCGCCCGCCGCGCCATGCTGTGCGCGGCTCGAGAGTGGCGCGCGC 139
 Db 1 TGCCTCCGTCCTCCCGCCCGCCGCGCCATGCTGTGCGCGGCTCGAGAGTGGCGCGCGC 60

QY 140 CTTTGCCCTCCCGCCGACAGGAGCGGAGCGCGCGCTCCGCGCGCACGGGAGCTGC 199
 Db 61 CTTTGCCCTCCCGCCGACAGGAGCGGAGCGCGCGCTCCGCGCGCACGGGAGCTGC 120

QY 200 AGTACCTGGGGGAGATCCACATCTCCGCTCGGCGCTCAGGAAGGACGACCGCAGG 259
 Db 121 AGTACCTGGGGGAGATCCACATCTCCGCTCGGCGCTCAGGAAGGACGACCGCAGG 180

QY 260 GCACCGGACCCCTGTGGTATTTCCGATGCGAGGCGGCTACAGCTTGAGATGAATTC 319
 Db 181 GCACCGGACCCCTGTGGTATTTCCGATGCGAGGCGGCTACAGCTTGAGATGAATTC 240

QY 320 CTCTGCTGCAACCAACAGCTGTCTTCTGGAAGGGTGTCTTGGAGGAGTGTCTGTGGTTTA 379
 Db 241 CTCTGCTGCAACCAACAGCTGTCTTCTGGAAGGGTGTCTTGGAGGAGTGTCTGTGGTTTA 300

QY 380 TCAAGGATCCCAAAATGCTAAAGAGCTGTCTTCCAAAGGAGTGAATAATCTGGGATGCCA 439
 Db 301 TCAAGGATCCCAAAATGCTAAAGAGCTGTCTTCCAAAGGAGTGAATAATCTGGGATGCCA 360

QY 440 ATGGATCCCGAGACATTTTGGAGACAGCTTGGATCTCCACAGAGAAGAGGGGACTTGG 499
 Db 361 ATGGATCCCGAGACATTTTGGAGACAGCTTGGATCTCCACAGAGAAGAGGGGACTTGG 420

QY 500 GCCCAGTTTATGCTTCCAGTGGAGGATTTTGGGCGCAATACACAGATATGAATCAG 559
 Db 421 GCCCAGTTTATGCTTCCAGTGGAGGATTTTGGGCGCAATACACAGATATGAATCAG 480

QY 560 ATTATTCAGAGACAGGAGTGTGACCAACTGCAAAAGAGTGTGACACCACTCAAAACCAACC 619
 Db 481 ATTATTCAGAGACAGGAGTGTGACCAACTGCAAAAGAGTGTGACACCACTCAAAACCAACC 540

QY 620 CTGACGACAGAAATCATGTGCGCTTGGAAATCCAAAGAGATCTTCTCTGATGGCGC 679
 Db 541 CTGACGACAGAAATCATGTGCGCTTGGAAATCCAAAGAGATCTTCTCTGATGGCGC 600

QY 680 TGCCTCCATGCCATGCCCTCTGCCAGTTCATGTGTTGAACAGTGTCTCTGCGCAGC 739
 Db 601 TGCCTCCATGCCATGCCCTCTGCCAGTTCATGTGTTGAACAGTGTCTCTGCGCAGC 660

QY 740 TGTACACAGATCGGAGACATGGGCTCGGTGTGCTTCAACATCCGAGCTACGCCCC 799
 Db 661 TGTACACAGATCGGAGACATGGGCTCGGTGTGCTTCAACATCCGAGCTACGCCCC 720

QY 800 TGCTCAGGTACATGATGCGCACATCAGGCGCTTGAAGCCAGTGTGACTTTATATACACTT 859
 Db 721 TGCTCAGGTACATGATGCGCACATCAGGCGCTTGAAGCCAGTGTGACTTTATATACACTT 780

QY 860 T-GGGAGATGCATATTTTACCTGAATCACAATCGAGCCACTGAAATTCAGCTTCAGCGA 918
 Db 781 TGGGAGATGCATATTTTACCTGAATCACAATCGAGCCACTGAAATTCAGCTTCAGCGA 840

QY 919 GAACCCAGACTTTCCTCCAAAGCTCAGGATTTCTGAAAAGTTTGAGAAAATTCAGACTTC 978
 Db 841 GAACCCAGACTTTCCTCCAAAGCTCAGGATTTCTGAAAAGTTTGAGAAAATTCAGACTTC 900

QY 979 AAAGCTGAAGACTTTCAGATTCAGAGGTACATCCGATCCCAACTATTAAATGGAATG 1038
 Db 901 AAAGCTGAAGACTTTCAGATTCAGAGGTACATTCGATCCCAACTATTAAATGGAATG 958

QY 1039 GCTGTTAGGGTGTCTTCAA 1058
 Db 959 GCTGTTAGGGTGTCTTCAA 978

RESULT 7
 AL542409
 LOCUS AL542409 LRI FL002 PL1 Homo sapiens cDNA clone CS0DE010YH04 5 prime
 DEFINITION , mRNA sequence.
 ACCESSION AL542409
 VERSION AL542409.1 GI:12874425
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 935)
 Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 CONTACT: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES		Location/Qualifiers	
source		1..935	
		/organism="Homo sapiens"	
		/db_xref="taxon:9606"	
		/clone="CS0DE010YH04"	
		/clone_lib="LTI_FL002_Pfl1"	
		/lab_host="DH10B"	
		/note="Organ: placenta; Vector: pCMWSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMWSPORT 6 vector. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville , Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"	
BASE COUNT	208 a	279 c	254 g 190 t 4 others
RIGIN			
	Query·Match	60.7%;	Score 933; DB 9; Length 935;
	Best Local Similarity	99.8%;	Pred.No. 3.4e-232;
	Matches 931; Conservative	4;	Mismatches 0; Indels 0; Gaps 0;
y	30	CCTCCCGCGCGGCACATTGGCTTCGCTCCGCCGCGGCACATTCGCTGCGCTCCGTC	89
b	1	CCTCCCGCGCGGCACATTGGCTTCGCTCCGCCGCGGCACATTCGCTGCGCTCCGTC	60
y	90	CCCCCGCCGCGCGCCCATGTCTTGCGCGGCTCGGAGTCTCGCGCGCGGCGCCTTGCGCCC	149
b	61	CCCCCGCCGCGCGCCCATGTCTTGCGCGGCTCGGAGTCTCGCGCGCGGCGCCTTGCGCCC	120
y	150	CGCGCGCACAGGAGCGGAGCGCGAGCGCGTTCGCCCGCACCGGGAGCTGCAGTA	209
b	121	CGCGCGCACAGGAGCGGAGCGCGAGCGCGTTCGCCCGCACCGGGAGCTGCAGTA	180
y	210	GCAGATCCAAACACATCTTCGCTGCGGCGTCAGGAGACACACCGCACGGGACCGGCAC	269
b	181	GCAGATCCAAACACATCTTCGCTGCGGCGTCAGGAAGACACCGCACGGGACCGGCAC	240
y	270	CTGTGCGTATTTCGGCATGCAGCGCGCTACAGCTGAGAGATGAATTCCTCTGCTGAC	329
b	241	CTGTGCGTATTTCGGCATGCAGCGCGCTACAGCTGARATGAATTCCTCTGCTGAC	300
y	330	AACCAACGCTGTGTTCTGGAAGGCTGTTTTGAGAGCTTGCTGTGTTTTATCAAGGATC	389
b	301	AACCAACGCTGTGTTCTGGAAGGCTGTTTTGAGAGCTTGCTGTGTTTTATCAAGGATC	360
y	390	CACAATGCTAAAGAGCTGTTCTTCCAAGGGAGTGAATAATTCGGGATGCCAATGGATCCCG	449
b	361	CACAATGCTAAARAGTGTCTTCCAAGGGAGTGAATAATTCGGGATGCCAATGGATCCCG	420
y	450	AGACTTTTGGACAGCGCTGGATTTCCACAGAGAAGAGGGAATTGGGGCCCGATTTA	509
b	421	AGACTTTTGGACAGCGCTGGATTTCCACAGAGAAGAGGGAATTGGGGCCCGATTTA	480
y	510	TGGCTTCCAGTGGAGGCAATTTTGGGSCAGAAATACAGAGATATGAATCAGATTATTCAGG	569
b	481	TGGCTTCCAGTGGAGGCAATTTTGGGSCAGAAATACAGAGATATGAATCAGATTATTCAGG	540
y	570	ACAGGGAGTTCACCACCTGCAAGAGTGAATTCACACCATCAAACCAACCTTCAGCAGAG	629
b	541	ACAGGGAGTTCACCACCTGCAAGAGTGAATTCACACCATCAAACCAACCTTCAGCAGAG	600
y	630	AAGAATCATCATGTGCGCTTGGAAATCCAAAGATCTTCTCTGATGGCGCTGCCTCCATG	689
b	601	AAGAATCATCATGTGCGCTTGGAAATCCAAAGATCTTCTCTGATGGCGCTGCCTCCATG	660
y	690	CCATGCCCTCTGCGAGTTCTATGTGTTGAACAGTGTCTCTGCGAGCTGTATTCAGAG	749
b	661	CCATGCCCTCTGCGAGTTCTATGTGTTGAACAGTGTCTCTGCGAGCTGTATTCAGAG	720
	750	ATCGGGAGACATGGGCGCTCGGTGTGCTTTTCAACATGCCAGCTAGCGCCCTGCTCACGTA	809

Db	721	ATCGGGAGACATGGGCTCGGTGTGCCTTTTCACATGCGCAGCTAGCCCTGCTCAGCTA	780
Qy	810	CATGATTGGCCACATCAACGGGCTTAAGCCAGGTGATTTTATACACTTTTGGAGATGC	869
Db	781	CATGATTGGCCACATCAACGGGCTTAAGCCAGGTGATTTTATACACTTTTGGAGATGC	840
Qy	870	ACATATTACTGAATCAATCATGAGCCACTGAAATTTACAGTTTACGCGAGAACCCAGACC	929
Db	841	ACATATTACTGAATCAATCATGAGCCACTGAAATTTACAGTTTACGCGAGAACCCAGACC	900
Qy	930	TTTCCCAAAGCTCAGGATTCITCGAAAGTTGAGA	964
Db	901	TTTCCCAAAGCTCAGGATTCITCGAAAGTTGAGA	935
RESULT 8			
AL551990			
LOCUS			
DEFINITION			
AL551990 LFI_NFL006.PL2 Homo sapiens cDNA clone CS0DI060YE04 5			
prime, mRNA sequence.			
ACCESSION			
AL551990			
VERSION			
AL551990.1 GI:12890459			
KEYWORDS			
EST.			
SOURCE			
Homo sapiens			
ORGANISM			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE			
1. (bases 1 to 938)			
AUTHORS			
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.			
TITLE			
Full-length cDNA libraries and normalization			
JOURNAL			
Unpublished (2001)			
COMMENT			
Contact: Genoscope			
Genoscope - Centre National de Sequencage			
BP 191 91006 EVRY cedex - France			
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.			
FEATURES			
source			
1..938			
/organism="Homo sapiens"			
/db_xref="taxon:9606"			
/clone="CS0DI060YE04"			
/tissue_lib="LFI_NFL006.PL2"			
/tissue_type="placenta"			
/note="Vector: pCMVSPORT 6; Site1: NotI; 1st strand cDNA			
was primed with a NotI-oligo(dT) primer. Five prime end			
enriched, double-stranded cDNA was digested with Not I and			
cloned into the Not I and Eco RV sites of the pCMVSPORT 6			
vector. Library was normalized. Library was constructed by			
Life Technologies. Contact : Feng Liang Life Technologies,			
a division of Invitrogen 9800 Medical Center Drive			
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371			
Email : fliang@lifestech.com URL :			
http://fulllength.invitrogen.com"			
BASE COUNT			
ORIGIN			
202 a 288 c 259 g 188 t 1 others			
Query Match			
60.7%; Score 931.6; DB 9; Length 938;			
Best Local Similarity 99.9%; Pred. No. 7.8e-232;			
Matches 931; Conservative 1; Mismatches 0; Indels 0; Gaps 0;			
Qy	14	CCACTTGGCTGTGCTCGTCCCGCGGCCACTTGGCGCTCGCTCCGTCGCCCGCGCCAC 73	
Db	7	CCACTTGGCTGTGCTCGTCCCGCGGCCACTTGGCGCTCGCTCCGTCGCCCGCGCCAC 66	
Qy	74	TTTGCGCTGCTTCGCTCCCGCGGCCACTTGGCGCTCGCTCGCGCTCGAGTGCCTCG 133	
Db	67	TTTGCGCTGCTTCGCTCCCGCGGCCACTTGGCGCTCGCTCGCGCTCGAGTGCCTCG 126	
Qy	134	GCGGCGCTTCGCTCCCGCGGCCACTTGGCGCTCGCTCGCGCTCGAGTGCCTCG 193	
Db	127	GCGGCGCTTCGCTCCCGCGGCCACTTGGCGCTCGCTCGCGCTCGAGTGCCTCG 186	
Qy	194	AGCTCGAGTACCTGGGGCAGATCCCAACATCTCCGCTGCGCGCGTCAAGAACCGAAC 253	

```

187 AGCTGACGTACTCTGGGACATCAACATCTCCGCTGCGGCTCAGAGACACC 246
254 GCACGGGACCGGACACCTCTGCGTATTCGGCATCAGGCGCGCTACAGCCTCAGAGATG 313
247 GCACGGGACCGGACACCTCTGCGTATTCGGCATCAGGCGCGCTACAGCCTCAGAGATG 306
314 AATTCCTCTGCTGCACACCAACAGTGTCTTCTGGAAGGGTGTTCGAGAGATTCCTCT 373
307 AATTCCTCTGCTGCACACCAACAGTGTCTTCTGGAAGGGTGTTCGAGAGATTCCTCT 366
374 GGTTCATCAAGGGATCCCAAAATGCTAAAGAGTGTCTTCAAGGGATGTAATTCGGG 433
367 GGTTCATCAAGGGATCCCAAAATGCTAAAGAGTGTCTTCAAGGGATGTAATTCGGG 426
434 ATGCCAATGGATCCCGAGACTTTTGGACAGCTGGATTCACACAGAGAGAGGGG 493
427 ATGCCAATGGATCCCGAGACTTTTGGACAGCTGGATTCACACAGAGAGAGGGG 486
494 ACTTGGGCCCAGTTTATGGCTTCAGTGGAGGCAATTTGGGCGAGATACAGAGATATGG 553
487 ACTTGGGCCCAGTTTATGGCTTCAGTGGAGGCAATTTTGGGCGAGATACAGAGATATGG 546
554 AATCAGATTTTTCAGGACAGGAGTTCACCACTGCAAGAGTGTATTCACACCATCAAAA 613
547 AATCAGATTTTTCAGGACAGGAGTTCACCACTGCAAGAGTGTATTCACACCATCAAAA 606
614 CCAACCTTCAGACAGAGAGATCATCATGTGCGCTTGAATCCAGAGATCTTCCTCTGA 673
607 CCAACCTTCAGACAGAGAGATCATCATGTGCGCTTGAATCCAGAGATCTTCCTCTGA 666
674 TGGCGTGTCTCCATGCCATCCCTCTGCCAGTTCTATGTGTGTAACAGTGTGTCTCT 733
667 TGGCGTGTCTCCATGCCATCCCTCTGCCAGTTCTATGTGTGTAACAGTGTGTCTCT 726
734 GCCAGTGTACAGAGATCGGGAGACATGGGCTCGGTGTGCTTCAACATGCCAGCT 793
727 GCCAGTGTACAGAGATCGGGAGACATGGGCTCGGTGTGCTTCAACATGCCAGCT 786
794 ACGCCTGTCTACGTACATGATTCGGACATCAGGCGCTGAAGCCAGTGTATATAC 853
787 ACGCCTGTCTACGTACATGATTCGGACATCAGGCGCTGAAGCCAGTGTATATAC 846
854 ACATTTGGAGATGCACATTTTACCTGAATCAGATCAGCCTGAAATTCAGCTTC 913
847 ACATTTGGAGATGCACATTTTACCTGAATCAGATCAGCCTGAAATTCAGCTTC 906
914 AGCGAGACCCAGACCTTTCCCAAGCTCAGG 945
907 AGCGAGACCCAGACCTTTCCCAAGCTCAGG 938

```

```

RESULT 9
AL515911 995 bp mRNA linear EST 13-FEB-2001
LOCUS AL515911 LTI_NFL011_NBC1 Homo sapiens cDNA clone CSODA001YD20 5
DEFINITION prime, mRNA sequence.
ACCESSION AL515911
VERSION AL515911.1 GI:12779404
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 995)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
          Genoscope - Centre National de Sequencage
          BP 191 91006 EVRY cedex - France
          Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
          Location/Qualifiers
FEATURES

```

```

1. .995
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CSODA001YD20"
/clone_lib="LTI_NFL011_NBC1"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library is not normalized, but is the control for
the normalized libraries. Library was constructed by Life
Technologies. Contact : Feng Liang Life Technologies, a
division of Invitrogen 9800 Medical Center Drive Rockville
Maryland 20850, USA Fax : (1) 301 610 8371 Email :
fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT 239 a 272 c 266 g 209 t 9 others
ORIGIN
Query Match 60.2%; Score 924; DB 9; Length 995;
Best Local Similarity 98.1%; Pred. No. 7.6e-230;
Matches 977; Conservative 9; Mismatches 5; Indels 5; Gaps 5;
QY 53 GCGCGGCGACATTCGCTCCGTCCTCCGCGCGCGCGCGCGCGCGCGCGCGCTC 122
Db 1 GCGCGGCGACATTCGCTCCGTCCTCCGCGCGCGCGCGCGCGCGCGCGCTC 60
QY 123 GGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTC 182
Db 61 GGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTC 120
QY 183 GCCGCAAGGGGAGTGCAGTACCTGGGCGAGATCCAAACATCTCCGCTCGCGCGTCA 242
Db 121 GCCGCAAGGGGAGTGCAGTACCTGGGCGAGATCCAAACATCTCCGCTCGCGCGTCA 180
QY 243 GAAGAGACCGGACCGGCGACCGGCGACCGTGTGCGTATTCGGCATTCAGGCGCGCTCA 302
Db 181 GAAGAGACCGGACCGGCGACCGGCGACCGTGTGCGTATTCGGCATTCAGGCGCGCTCA 240
QY 303 CTTGAGAGATGAATTCCTCTCTGCAACAACAACTGTGTCTCGAGAGGTGTTTGA 362
Db 241 CCTGARATATGAATTCCTCTCTGCAACAACAACTGTGTCTCGAGAGGTGTTTGA 300
QY 363 GGAGTCTGTGTGTATTCAGGGGATCCAAATGCTAAAGAGCTGTCTTCCAAAGGAGT 422
Db 301 GGAGTCTGTGTGTATCAAGGGATCCAAATGCTAAAGAGCTGTCTTCCAAAGGAGT 360
QY 423 GAAATCTGGGATGCCAATGGATCCGAGACTTTTGGACAGCTTCTCCACCAG 482
Db 361 AATAATCTGGGATGCCAATGGATCCGAGACTTTTGGACAGCTTCTCCACCAG 420
QY 483 AGAAGAGGGGACTTTGGGCGGCGGCTTATGGCTTCCAGTGGAGGCAATTTGGGGCAGAATA 542
Db 421 AGAAGAGGGGACTTTGGGCGGCGGCTTATGGCTTCCAGTGGAGGCAATTTGGGGCAGAATA 480
QY 543 CAGAGATATGGAATCAGATTATTCAGGACAGGAGTTCACCACTGCCAAGAGTGAATGA 602
Db 481 CAGAGATATGGAATCAGATTATTCAGGACAGGAGTTCACCACTGCCAAGAGTGAATGA 540
QY 603 CACCATCAAAACCAACCTTCAGGACAGAGATCATCATGTGCGTTCGAATCCAAAGAGA 662
Db 541 CACCATCAAAACCAACCTTCAGGACAGAGATCATCATGTGCGTTCGAATCCAAAGAGA 600
QY 663 TCTTCTCTGATGGCGCTCCCTCCATGCGATCCGCTCTGCAAGTTCATGTGTGTAACAG 722
Db 601 TCTTCTCTGATGGCGCTCCCTCCATGCGATCCGCTCTGCAAGTTCATGTGTGTAACAG 660
QY 723 TGAGCTGTCTGCGACGCTGTACAGAGATCGGGAGAGATGGGCGCTCGGTGTGCTTTCAA 782
Db 661 TGAGCTGTCTGCGACGCTGTACAGAGATCGGGAGAGATGGGCGCTCGGTGTGCTTTCAA 720

```

source


```

Qy 633 AATCATCATGTCGCTGGGATCAAGAGATCTTCTCTGATGGCGCTGCTCCATGCA 692
Db 601 AATCATCATGTCGCTGGGATCAAGAGATCTTCTCTGATGGCGCTGCTCCATGCA 660
Qy 693 TGGCCCTCTGCGCATGTTCTATGTTGATGACAGTGTCTCTGATGGCGCTGCTCCATGCA 752
Db 661 TGGCCCTCTGCGCATGTTCTATGTTGATGACAGTGTCTCTGATGGCGCTGCTCCATGCA 720
Qy 753 GGGAGACATGGGCTCGCTGGGATCAAGAGATCTTCTCTGATGGCGCTGCTCCATGCA 812
Db 721 GGGAGACATGGGCTCGCTGGGATCAAGAGATCTTCTCTGATGGCGCTGCTCCATGCA 780
Qy 813 GATTGGGACATCAAGGCTCGGCTGGGATCAAGAGATCTTCTCTGATGGCGCTGCTCCATGCA 870
Db 781 GATTGGGACATCAAGGCTCGGCTGGGATCAAGAGATCTTCTCTGATGGCGCTGCTCCATGCA 840
Qy 871 CATATTACTTCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 927
Db 841 CATATTACTTCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 900
Qy 928 CTTTT-CCCAAGCTCAGGATCTT---CGAAAAGTGGAAAATGATGA-CITCAAG 982
Db 901 CTTTTCCCAAGCTCAGGATCTT---CGAAAAGTGGAAAATGATGA-CITCAAG 960
Qy 983 CTCGAACATTT--CAGATTGAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1040
Db 961 CTGAAAATTTTTCAGAAATGAGGACATCCCGCATCATCTTTTAAATGGAATGGG 1020

Qy 1041 T 1041
Db 1021 T 1021

```

RESULT 14
AL515910/c
LOCUS
DEFINITION
AL515910 LTI_NFL011_NBC1 Homo sapiens cDNA clone CS0DA001YD20 3
prime, mRNA sequence.

ACCESSION
AL515910
VERSION
AL515910.1 GI:12779403
KEYWORDS
EST.
SOURCE
human.

ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)

CONTACT
Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source
1. .923
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DA001YD20"
/clone_lib="LTI_NFL011_NBC1"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library is not normalized, but is the control for
the normalized libraries. Library was constructed by Life
Technologies. Contact : Peng Liang Life Technologies, a
division of Invitrogen 9800 Medical Center Drive Rockville
Maryland 20850, USA Fax : (1) 301 610 8371 Email :
fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

```

BASE COUNT      249 a      206 c      188 g      269 t      11 others
ORIGIN
Query Match      57.8%; Score 888; DB 9; Length 923;
Best Local Similarity 97.4%; Pred. No. 1.8e-220;
Matches 899; Conservative 10; Mismatches 13; Indels 1; Gaps 1;
Qy 570 ACAGGAGTTGACCAACTGCAAGAGTGTGACACATCAAAACAA-CCCTGACGACA 628
Db 923 ACAGGAGTTGACCAACTGCAAGAGTGTGACACATCAAAACAA-CCCTGACGACA 864
Qy 629 GAAGATCATCATGCTGGCTGGATCAAGAGATCTTCTCTGATGGCGCTGCTCCAT 688
Db 863 GAAGATCATCATGCTGGCTGGATCAAGAGATCTTCTCTGATGGCGCTGCTCCAT 804
Qy 689 GCCATGCCCTGCGCAGTTCATGTGTGAACAGTGTGCTGCGCAGCTGTACACAGA 748
Db 803 GCCATGCCCTGCGCAGTTCATGTGTGAACAGTGTGCTGCGCAGCTGTACACAGA 744
Qy 749 GATCGGAGACATGGGCTCGCTGCTTTCAACATCGCCAGCTACGCCCTGCTCAGCT 808
Db 743 GATCGGAGACATGGGCTCGCTGCTTTCAACATCGCCAGCTACGCCCTGCTCAGCT 684
Qy 809 ACATGATTGGGACATCAAGGCTGGAAGCCAGGTGACTTTATACACACTTTGGGAGATG 868
Db 683 ACATGATTGGGACATCAAGGCTGGAAGCCAGGTGACTTTAAACACACTTTGGGAGATG 624
Qy 869 CACATATTACCTGAATCAATCGAGCCACTGAAAATTCAGCTTCAGCGAGAACCCAGAC 928
Db 623 CACATATTACCTGAATCAATCGAGCCACTGAAAATTCAGCTTCAGCGAGAACCCAGAC 564
Qy 929 CTTTCCCAAGCTCAGGATCTTTCGAAAAGTGTGAAAATTCAGCTTCAGCTTCAGCTCAAG 988
Db 563 CTTTCCCAAGCTCAGGATCTTTCGAAAAGTGTGAAAATTCAGCTTCAGCTTCAGCTCAAG 504
Qy 989 ACTTTCAGATTGAAGGTGACAAATCGCATCCAACTTAAATGAAATGCTGTTAGG 1048
Db 503 ACTTTCAGATTGAAGGTGACAAATCGCATCCAACTTAAATGAAATGCTGTTAGG 444
Qy 1049 GTGCTTTCAAAAGGAGTGAAGGATTTGTGCTTTAGGGGTTGGGCTGGATGCCGAG 1108
Db 443 GTGCTTTCAAAAGGAGTGAAGGATTTGTGCTTTAGGGGTTGGGCTGGATGCCGAG 384
Qy 1109 GTAAAGTTCTTTTCTCTCTAAAGAAAGAACTAGAGTCAAAAATCTGCTCGTGAACCT 1168
Db 383 GTAAAGTTCTTTTCTCTCTAAAGAAAGAACTAGAGTCAAAAATCTGCTCGTGAACCT 324
Qy 1169 ATCAGTTATTAAATTTTAAAGGATGTGCCACTGGGAAATGTAACTGTGCCAGTTCTTCC 1228
Db 323 WTCAGTTATTAAATTTTAAAGGATGTGCCACTGGGAAATGTAACTGTGCCAGTTCTTCC 264
Qy 1229 ATAAATAAAGGCTTTCAGTTTACCTCAGCTGAGGTTCTGACAAATCTGAGGTTTGAACA 1288
Db 263 ATAAATAAAGGCTTTCAGTTTACCTCAGCTGAGGTTTCTGACAAATCTGAGGTTTGAACA 204
Qy 1289 AAGTGAGGAGATCAAAATCTATGTCTCTAGCAAAAACATGTATGTGCTTCAATCCC 1348
Db 203 AAGTGAGGAGATCAAAATCTATGTCTCTAGCAAAAACATGTATGTGCTTCAATCCC 144
Qy 1349 ACGTACTTATAAAGAGGTTGGTGAATTTCAAGCTTATTTTGAATATTTTAGAATA 1408
Db 143 ACGTACTTATAAAGAGGTTGGTGAATTTCAAGCTTATTTTGAATATTTTAGAATA 84
Qy 1409 TTTTAAAGATTTCAAGCTTATTTCCCTCAATCTGAGGAGGAGTGTGATACACCTTCATC 1468
Db 83 TTTTAAAGATTTCAAGCTTATTTCCCTCAATCTGAGGAGGAGTGTGATACACCTTCATC 24
Qy 1469 ATGATGTAGAGTGTGTTATGAA 1491
Db 23 ATGATGTAGAGTGTGTTATGAA 1

```

RESULT 15

AL578646/c
LOCUS
AL578646 LTI NFL006 PL2 Homo sapiens cDNA clone CSODK004YL08 3
prime, mRNA sequence.
AL578646
AL578646.1 GI:12942922
EST.
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1. (bases 1 to 906)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
1..906
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CSODK004YL08"
/clone_lib="LTI NFL006 PL2"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Peng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
ASE COUNT 256 a 200 c 183 g 264 t 3 others
ORIGIN
Query Match 57.8%; Score 887.6; DB 9; Length 906;
Best Local Similarity 99.0%; Pred. No. 2.3e-220;
Matches 898; Conservative 3; Mismatches 5; Indels 1; Gaps 1;
605 CCATCAAAACCAACCTGACGACAGAGAAATCATATGTGGCTTGGATCCAGAGATC 664
b 906 CCATCAAAACCAACCTGACGACAGAGAAATCATATGTGGCTTGGATCCAGAGATC 847
665 TTCCTCTGATGGCGCTGCTTCATGCCAGCCCTCTGCCAGTTCATATGTGGTGAACAGTG 724
b 846 TTCCTCTGATGGCGCTGCTTCATGCCAGCCCTCTGCCAGTTCATATGT-GTGAACAGTG 788
725 AGCTGTCTGCGCAGCTGTACAGAGATCGGAGACATGGGCTCGGTGCGCTTCAACA 784
b 787 AGCTGTCTGCGCAGCTGTACAGAGATCGGAGACATGGGCTCGGTGCGCTTCAACA 728
785 TCGCCAGGTGAGCCCTGCTCAGCTACATGATTCGCGACATCAACGGGCTGAGCCAGGTG 844
b 727 TCGCCCAACGAGCCCTGCTCAGCTACATGATTCGCGACATCAACGGGCTGAGCCAGGTG 668
845 ACTTTATACACACTTTGGGAGATGCATATTTTACCTGAATTCACATCGAGCCACTGAAA 904
b 667 ACTTTATACACACTTTGGGAGATGCATATTTTACCTGAATTCACATCGAGCCACTGAAA 608
905 TTCAGCTTCAGCGAGAACCCAGACCTTCCCAAAGCTCAGGATTCCTCGAAAAGTTGAGA 964
b 607 TTCAGCTTCAGCGAGAACCCAGACCTTCCCAAAGCTCAGGATTCCTCGAAAAGTTGAGA 548
965 AAATTGATGACTTCAAAGCTGAAGACTTTTCAGATTCAGAGGGTACAAATCCGATCCAACTA 1024
b 547 AAATTGATGACTTCAAAGCTGAAGACTTTTCAGATTCAGAGGGTACAAATCCGATCCAACTA 488
1025 TTAATAATGGAATGGCTCTTTAGGGTCTCTTCAAAGGAGCTNGAAGGATATTGTCACTCT 1084

Db 487 TTAATAATGGAATGGCTCTTTAGGGTCTCTTCAAAGGAGCTTGAAGGATATTGTCACTCT 428
QY 1085 TTAGGGGTTGGGCTGATGCCGAGTAAAGTCTCTTTTGGCTCTTAAAGAGAGGAACT 1144
Db 427 TTAGGGGTTGGGCTGATGCCGAGTAAAGTCTCTTTTGGCTCTTAAAGAGAGGAACT 368
QY 1145 AGGTCAAAATCTGTCCGAGACCTATCATCTTAAATTTTAAAGGATGTGCCCACTGGCA 1204
Db 367 AGGTCAAAATCTGTCCGAGACCTATCATCTTAAATTTTAAAGGATGTGCCCACTGGCA 308
QY 1205 AATCTAACTGTGCCAGTTCTTTTCCATAATAAAGGCTTTGAGTTAACTCACTGAGGGTAT 1264
Db 307 AATCTAACTGTGCCAGTTCTTTTCCATAATAAAGGCTTTGAGTTAACTCACTGAGGGTAT 248
QY 1265 CTGACAAATGCTGAGGTTATGAACAAACTGAGGAGAAATGAATGTATGTCTTTAGCAAA 1324
Db 247 CTGACAAATGCTGAGGTTATGAACAAAGTGAAGGAGAAATGAATGTATGTCTTTAGCAAA 188
QY 1325 AACATGTATGTGCATTTCAATCCCACTGATCTTATAAAGAGGTTGGTGAATTTCAACAAGC 1384
Db 187 AACATGTATGTGCATTTCAATCCCACTGATCTTATAAAGAGGTTGGTGAATTTCAACAAGC 128
QY 1385 TATTTTGAATATTTTGTAGATATTTTAAAGATTTTCAAGCTATTCCTCAATCTGA 1444
Db 127 TATTTTGAATATTTTGTAGATATTTTAAAGATTTTCAAGCTATTCCTCAATCTGA 68
QY 1445 GGGAGCTGAGTAACCACTGATCATGATGTAGAGTGTGTTTATCAACTTTANAGTTGTT 1504
Db 67 GGGAGCTGAGTAACCACTGATCATGATGTAGAGTGTGTTTATCAACTTTANAGTTGTT 8
QY 1505 TTATATG 1511
Db 7 TTATATG 1

Search completed: November 28, 2002, 22:40:04

Job time : 2033.8 secs


```

; ; TITLE OF INVENTION: Variants of alternative splicing
; ; FILE REFERENCE: 129181.4 Compugen
; ; CURRENT APPLICATION NUMBER: US/09/724,676
; ; CURRENT FILING DATE: 2000-11-28
; ; NUMBER OF SEQ ID NOS: 97222
; ; SOFTWARE: Patentin version 3.2
; ; SEQ ID NO 34515
; ; LENGTH: 2042
; ; TYPE: DNA
; ; ORGANISM: Homo sapiens
US-09-724-676-34515

```

```
Query Match          66.7%; Score 1024; DB 5; Length 2042;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1124; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

Qy 1 GGGGGGGGGGACCACTTTGGCTCGCTCCGTCGGCGGCCACTTGGCTGCCCTCGT 60

Db 2042 GGGGGGGGGGGACCACTTGGCTCGCTCCGTCGGCGGCCACTTGGCTGCCCTCGT 1883

61 CCGGCGGCGCCACTTCCCTCGCTCCGTCCCCCGCGCGGCATGCGCTGTGGCCGGC 120

db 1982 CCGCGCGCCACTTCGGCTGCCCTCCCGCCCGCCATGCCCTGTGCCGGC 1923

121 TCGGAGCTGCGCGCCGGCCCTTGCCCCCGCCGCACAGGAGCGGACGCCGAGCCGCGT 180

db 1922 TCGGAGCTGCGCGGCGCGGCGCTTCCCGCGCGCACAGGAGCGGGACGCCAGCCGCGT 1863

181 CCGCCGACGGGGAGCTGCAGTACCTGGGGCAGATCCAAACACATCCTCCGCTGGGGCGTC 240

1862 CCGCCGACGGGGAGCTGCAGTACCTGGGGCAGATCCAAACACATCCTCGCTGGGGGTC 1803

341 **ACGACGACCGACCGGCA**CCCTGTGGTATT**CGGCATGCAGGCGCGCTAC** 300

nt 1803 ACCACCAAGGCGACCGGCACCTGTCCGTTATTCGGCATGCGGCCGGCTAC 1743

360

1683

CGTCTTCGACGGGA 420

1623

480

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

[illegible]

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

[illegible][illegible][illegible]

QY
60I GACACCAH CAHHACCAHCCCCGAGCAGCACAGTACTAT CAA GAGGGCTTTT

D0 T442 GACACCATCAAAACCAACCCCTGGCGAGCAGATGATCATCATCTCCCTTCCCATTGCGAATGG

QY 661 GATCCTTCCTCGAAGGCGCCCTCCCATGGCAAGCCCTTCTATTGCGGTGTTTG
|||||

D6
1382 GATCTTCC^TTC^TGATGGGCGCTGCCCTCCATGCACATGCCCCCTCAGCCAGTTCTATTGTGGTGAACTCTCT

721 AGTGAGCTGTCTCGCCAGCTGTATCCAGAGATCGGGAGACATGGGCTTCGGGTGTCCTTC
QY

1322 AGTGAGCTGTCTCTGCCAGCTGTACACAGAGATCGGGAGACATGGGCCCTCGGTGTGCCCCCTTC 1280

781 AACATCGCCAGCTACGCCCTGCTCAGCTACATGATTCGGCACATCACGGGCTGAAGCCA 840

D_b 1262 AACATCGCCAGCTACGCCCTGCTCAGTACATGATTGGCACAATCACGGGCCCTGAAGCCA 1205

QY 841 GG TGACTTTATACACACTTTGGGAGATGCACATATTTACCTGAATCACATCGAGCCACTG 900

Db	1202	GGTGACTTTATACACACTTTGGGAGATGCACATATTTTACCTGAATCACATCGAGCACTG	1143
Qy	901	AAAATTGACGTTTCAGCGAGAACCCGAGACTCTTTCCAAAGACTCAGGATTCCTTCGAAAAGTT	960
Db	1142	AAAATTGACGTTTCAGCGAGAACCCGAGACTCTTTCCAAAGACTCAGGATTCCTTCGAAAAGTT	1083
Qy	961	GAGAAAATTGATGATCTTCAAAGCTGAAGACTTCCTCAGATTTGAAGGGTACAAATCCGCAATCCA	1020
Db	1082	GAGAAAATTGATGATCTTCAAAGCTGAAGACTTCCTCAGATTTGAAGGGTACAAATCCGCAATCCA	1023
Qy	1021	ACTATTAAATGGAATGGCTGTTTAGGGTGCCTTTCAAAGAGCTTNGAAGGATATTGTCA	1080
Db	1022	ACTATTAAATGGAATGGCTGTTTAGGGTGCCTTTCAAAGAGCTTNGAAGGATATTGTCA	963
Qy	1081	GTCTTTAGGGGTGGCGTCGATGCCGAGGTAAAAAGTTCTTTTTGCT	1126
Db	962	GTCTTTAGGGGTGGCGTCGATGCCGAGGTAAAAAGTTCTTTTTGCT	917

RESULT 6

US-09-724-676A-34515/c
: Sequence 34515. Application US/09724676A

```

1  / ORGANISM: Homo sapiens
2  /
3  / APPLICANT: CompuGen LTD
4  /
5  / TITLE OF INVENTION: Variants of alternative splicing
6  /
7  / FILE REFERENCE: 129181.4 CompuGen
8  /
9  / CURRENT APPLICATION NUMBER: US/09/724,676A
10 /
11 / CURRENT FILING DATE: 2000-11-28
12 /
13 / NUMBER OF SEQ ID NOS: 97222
14 /
15 / SOFTWARE: PatentIn version 3.2
16 /
17 / SEQ ID NO 34515
18 /
19 / LENGTH: 2042
20 /
21 / TYPE: DNA
22 /
23 / ORGANISM: Homo sapiens
24 /
25 / US-09-724-676A-34515

```

Query Match 66.7%; Score 1024; DB 5; Length 2042;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1124; Conservative 0; Mismatches 2; Indels 0

Qy	1	GGGGGGGGGGGACCATTTGGCTGCTTCCTGCTCCGGCGCGCCACATTTGAGCTGCTCCCTCGT	60
Db	2042	GGGGGGGGGGGACCATTTGGCTGCTTCCTGCTCCGGCGCGCCACATTTGAGCTGCTCCCTCGT	1983
Qy	61	CCGGCGGGCCACTTTCCGCTGCTCCGTCCTCCCGCGCGCGCCCATGCTCTGTGGCGCGGC	120
Db	1982	CCGGCGGGCCACTTTCCGCTGCTCCGTCCTCCCGCGCGCGCCCATGCTCTGTGGCGCGGC	1923
Qy	121	TGGGAGCTCCGGCGCGGGCCCTTTGCCCTCCCGCCGCA CAGGAGCGGGAAGCCGAGCGCGGT	180
Db	1922	TGGGAGCTCCGGCGCGGGCCCTTTGCCCTCCCGCCGCA CAGGAGCGGGAAGCCGAGCGCGGT	1863
Qy	181	CGCGCGCACGGGGAGCTGCAGTACTTGGGGCAGATCCAA CACATCCTCCGCTGCGGGCTC	240
Db	1862	CGCGCGCACGGGGAGCTGCAGTACTTGGGGCAGATCCAA CACATCCTCCGCTGCGGGCTC	1803
Qy	241	AGGAAGGACGACCGCACCGGCAACCTGTCGCTATTGCGCATG CAGCGCGCGCTAC	300
Db	1802	AGGAAGGACGACCGCACCGGCAACCTGTCGCTATTGCGCATG CAGCGCGCGCTAC	1743
Qy	301	AGCCTGAGAGATGAATTCCTCTGCTGACACAAAGTCGTGTTCTGGAAGGGTCTTTTG	360
Db	1742	AGCCTGAGAGATGAATTCCTCTGCTGACACAAAGTCGTGTTCTGGAAGGGTCTTTTG	1583
Qy	361	GAGGAGTTGCTGTGCTTTATCAAGGAGTCCCAAAATGCTAAAGAGCTGCTTCTCAAGGGA	420
Db	1682	GAGGAGTTGCTGTGCTTTATCAAGGAGTCCCAAAATGCTAAAGAGCTGCTTCTCAAGGGA	1623
Qy	421	GTGAAAATCTCGGATGCCAATGATCCGAGCTTTTGGACAGCCTGGGATTTCTCCACC	480
Db	1622	GTGAAAATCTCGGATGCCAATGATCCGAGCTTTTGGACAGCCTGGGATTTCTCCACC	1563

Db	148	GCCGAGCCGGCTCCGGCCGACGGGGAGCTGCAGTACCTCGGGGCGACATCCACACACATCCCTC	207
Qy	229	CGCTCGGGCGCTCAGGAAGGACGACCGCACCGGCACCGGCACCGCTCTGGGTATTTCGGCATG	288
Db	208	CGCTCGGGCGCTCAGGAAGGACGACCGCACCGGCACCGGCACCGCTCTGGGTATTTCGGCATG	267
Qy	289	CAGGCGCGCTACAGCCTTGAGAGATGAATTCCTCTGTCTGACAAACAAACGTTGTTCTTGG	348
Db	268	CAGGCGCGCTACAGCCTTGAGAGATGAATTCCTCTGTCTGACAAACAAACGTTGTTCTTGG	327
Qy	349	AAGGCTGTTTTGGAGGAGTTGCTGTGGTTATACAGGGATCCACAAATGCTAAAGAGCTG	408
Db	328	AAGGCTGTTTTGGAGGAGTTGCTGTGGTTATACAGGGATCCACAAATGCTAAAGAGCTG	387
Qy	409	TCTTCCAAAGGAGTGAAATCTGGGATGCCAATCGATCCCGAGACTTTTTCGACACGCTG	468
Db	388	TCTTCCAAAGGAGTGAAATCTGGGATGCCAATCGATCCCGAGACTTTTTCGACACGCTG	447
Qy	469	GGATTCTCCACAGAGAGAGGAGCTTGGGCCAGTTGAGTTCAGGATCCAGGAGATTCAGTGAGGCAT	528
Db	448	GGATTCTCCACAGAGAGAGGAGCTTGGGCCAGTTGAGTTCAGGATTCAGTGAGGCAT	507
Qy	529	TTTGGGCGAGATACAGAGATATGHAATCAGATTATTCAGGACAGGGAGTTGACAACTG	588
Db	508	TTTGGGCGAGATACAGAGATATGHAATCAGATTATTCAGGACAGGGAGTTGACCACTG	567
Qy	589	CAAGAAGTATTGACACATCAAAACCAACCTCGACACAGAGAATCATCATGTGCGCT	648
Db	568	CAAGAAGTATTGACACATCAAAACCAACCTCGACACAGAGAATCATCATGTGCGCT	627
Qy	649	TGGAATCCAAAGAGATCTTCCTGTGATGGCGCTGCTCCATGCGCATGCCCTCTGCCAGTTC	708
Db	628	TGGAATCCAAAGAGATCTTCCTGTGATGGCGCTGCTCCATGCGCATGCCCTCTGCCAGTTC	687
Qy	709	TATGTGGTGAAACAGTGAGCTGTCTCTCAGCTGTACACAGAGATCGGGAGACATGGGCTC	768
Db	688	TATGTGGTGAAACAGTGAGCTGTCTCTCAGCTGTACACAGAGATCGGGAGACATGGGCTC	747
Qy	769	GGTGTGCTTTCAATCGCCAGCTACGCCCTGCTCAGTACATGATTGCGCACATCAAG	828
Db	748	GGTGTGCTTTCAATCGCCAGCTACGCCCTGCTCAGTACATGATTGCGCACATCAAG	807
Qy	829	GGCTCGAGCCAGGTGACTTTATACACATTTGGGAGATGCACATATTTACCTGAATCAC	888
Db	808	GGCTCGAGCCAGGTGACTTTATACACATTTGGGAGATGCACATATTTACCTGAATCAC	867
Qy	889	ATCGAGCCACTGAAATTCAGCTTCAGCGAGAACCCAGACCTTCCCAAGCTCAGGATT	948
Db	868	ATCGAGCCACTGAAATTCAGCTTCAGCGAGAACCCAGACCTTCCCAAGCTCAGGATT	927
Qy	949	CTTCGAAAAGTTGAGAAATATGATGCTTCAAGCTGAGACCTTTCAGATTGAGGGTAC	1008
Db	928	CTTCGAAAAGTTGAGAAATATGATGCTTCAAGCTGAGACCTTTCAGATTGAGGGTAC	987
Qy	1009	AATCCGATCCAACTATTAAATGGAATGCGCTGTATAGGCTGCTTTCAAAGGAGCTNGA	1068
Db	988	AATCCGATCCAACTATTAAATGGAATGCGCTGTATAGGCTGCTTTCAAAGGAGCTCGA	1047
Qy	1069	AGGATATGTCAGTCTTTAGGGGTGGGCTGGATGCGAGGTAAAAAGTTCTTTTTGCTCT	1128
Db	1048	AGGATATGTCAGTCTTTAGGGGTGGGCTGGATGCGAGGTAAAAAGTTCTTTTTGCTCT	1107
Qy	1129	AAAAAANAAGGAACCTAGCTCAAAAATCTCTCCGTGACCTATCAGTTATTAATTTTAAAG	1188
Db	1108	AAAAAANAAGGAACCTAGCTCAAAAATCTCTCCGTGACCTATCAGTTATTAATTTTAAAG	1167
Qy	1189	GATGTGCTCCTGGCAAAATGTAATGTGCGAGTCTTTTCATTAATAAAGGCTTTGAGTT	1248
Db	1168	GATGTGCTCCTGGCAAAATGTAATGTGCGAGTCTTTTCATTAATAAAGGCTTTGAGTT	1227
Qy	1249	AACCTACTGAGGATATCTGACATGCTGAGGTATGAAACAAAGTTCAGGAGATGAATGT	1308
Db	1228	AACCTACTGAGGATATCTGACATGCTGAGGTATGAAACAAAGTTCAGGAGATGAATGT	1287

Qy	1309	ATGTGCTCTTAGAGAAAACATGTATCTGCATTTCAATCCACACTCTTATATAAGAAGGTT	1368
Db	1288	ATGTGCTCTTAGCAAAAACATGTATGTGCTTTCAATCCACACTCTTATAPAGAAGGTT	1347
Qy	1369	GTGTGAATTTTCCAAAGCTATTATTTTGGAAATATTTTGTAGAATATTTTACAAGCT	1428
Db	1348	GTGTGAATTTTCCAAAGCTATTATTTTGGAAATATTTTGTAGAATATTTTACAAGCT	1407
Qy	1429	ATTCCCTCAAATCTGAGGGAGCTGAGTAACACCAATCGATCATGCTAGAGTGTTGGTTAT	1488
Db	1408	ATTCCCTCAAATCTGAGGGAGCTGAGTAACACCAATCGATCATGCTAGAGTGTTGGTTAT	1467
Qy	1489	GAAGTTTA 1496	
Db	1468	GAAGTTTA 1475	

RESULT 3

```

RES001 3
US-09-724-676-34523/c
; Sequence 34523, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: CompuGen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 CompuGen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 34523
; LENGTH: 1942
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676-34523

```

Query Match	66.7%	Score 1024:	DB 5:	Length 1942;
-------------	-------	-------------	-------	--------------

Query Match	86.7%	Score Error, 25.5	Length Error, 25.5
Best Local Similarity	99.8%	Pred. No. 0;	
Matches 1124;	Conservative	0; Mismatches	2; Indels
		0; Gaps	0;

Qy	1	GGGGGGGGGGGACCACTTGGCTGCTCCGTCGCGCGCCACATTTGSCCTGCCTCGT	60
Db	1942	GGGGGGGGGGGACCACTTGGCTGCTCCGTCGCGCGCCACATTTGSCCTGCCTCGT	1883
Qy	61	CCCGCGCGCCACTTTCGCTTCCTCCGTCGCCCGCGCGCCATGCTGTGGCGGCG	120
Db	1882	CCCGCGCGCCACTTTCGCTTCCTCCGTCGCCCGCGCGCCATGCTGTGGCGGCG	1823
Qy	121	TCGAGAGTCGCGCGCGCCCTTTCGCCCGCGCACAGAGAGGGAAGCCGAGCGCGT	180
Db	1822	TCGAGAGTCGCGCGCGCCCTTTCGCCCGCGCACAGAGAGGGAAGCCGAGCGCGT	1763
Qy	181	CCCGCGCACGGGGAGCTGCAGTACTCTGGGGCAGATCCAAACACATCTCTCGCTGCGGCTC	240
Db	1762	CCCGCGCACGGGGAGCTGCAGTACTCTGGGGCAGATCCAAACACATCTCTCGCTGCGGCTC	1703
Qy	241	AGGAGGACGACCGCACGGGCACCGGCACCTCTGCGTATTTCGGCATGACGCGCGCTAC	300
Db	1702	AGGAGGACGACCGCACGGGCACCGGCACCTCTGCGTATTTCGGCATGACGCGCGCTAC	1643
Qy	301	AGCCTGAGAGATGAATTCCTCTGCTGCAACACCAAACGTGTGTCTCGAAGGGTCTTTTG	360
Db	1642	AGCCTGAGAGGTGAATTCCTCTGCTGCAACACCAAACGTGTGTCTCGAAGGGTCTTTTG	1583
Qy	361	GAGGAGTTCCTGTGGTTTATCAAGGGATCCAAATGCTAAAGAGCTGTCTTCCAAAGGA	420
Db	1582	GAGGAGTTCCTGTGGTTTATCAAGGGATCCAAATGCTAAAGAGCTGTCTTCCAAAGGA	1523
Qy	421	GTCAAAATCTGGATGCGCAATGATCCCGAGACTTTTGTGACAGCCTGGGATTCCTCAAC	480
Db	1522	GTCAAAATCTGGATGCGCAATGATCCCGAGACTTTTGTGACAGCCTGGGATTCCTCAAC	1463
Qy	481	AGAGAAAGAGGGGACTTGGGCCCGCATTTATGGCTTCAGTGGAGGCAATTTTGGGCGAGAA	540

Db	61	CCCGCGCGCCACTTGCCTTGCCTCCGTCGCCCGCGCGGCCCATGCTGTGTGGCGGC	120
QY	121	TCGGAGCTGCCGCGCGGCCCTTGGCCCCCGCCGCGCAGGAGCGGAGCGCGAGCCGGGT	180
Db	121	TCGGAGCTTCCCGCGCGGCCCTTGGCCCCCGCGCGCAGGAGCGGAGCGCGAGCCGGGT	180
QY	181	CCGCGCGCAGCGGGAGCTGCAGTACTCTGGGGCAGATCCAAACAATCCTCCGTTCGGCGTC	240
Db	181	CCGCGCGCAGCGGGAGCTGCAGTACTCTGGGGCAGATCCAAACAATCCTCCGTTCGGCGTC	240
QY	241	AGGAAGGACGACCGCAGCGGCAACCGGCAACCTGTCGGTATTCGGCATCCAGCGCGCTAC	300
Db	241	AGGAAGGACGACCGCAGCGGCAACCGGCAACCTGTCGGTATTCGGCATCCAGCGCGCTAC	300
QY	301	AGCCTGAGAGATGAATTCCTCTGTGACCAACAAACGTCGTCTGCAAGGGGTGTTTG	360
Db	301	AGCCTGAGAGATGAATTCCTCTGTGACCAACAAACGTCGTCTGCAAGGGGTGTTTG	360
QY	361	GAGGAGTGTGCTGCTTTATCAAGGATCCAAATCTTAAAGAGCTGCTTCCAAAGGA	420
Db	361	GAGGAGTGTGCTGCTTTATCAAGGATCCAAATCTTAAAGAGCTGCTTCCAAAGGA	420
QY	421	GTGAAAATCTGGGATGCAATGGATCCCGAGACTTTTGGCAGAGCTGGGATTCCTCAAC	480
Db	421	GTGAAAATCTGGGATGCAATGGATCCCGAGACTTTTGGCAGAGCTGGGATTCCTCAAC	480
QY	481	AGAGAAGAGGGGACTTGGGCCCGAGTTATGGCTTCAGGTGGAGGACATTTTGGGGCAGAA	540
Db	481	AGAGAAGAGGGGACTTGGGCCCGAGTTATGGCTTCAGGTGGAGGACATTTTGGGGCAGAA	540
QY	541	TACAGAGATATGGAATCAGATTTATTCAGGACAGGGAGTGCACCACTGCAAAAGAGTGATT	600
Db	541	TACAGAGATATGGAATCAGATTTATTCAGGACAGGGAGTGCACCACTGCAAAAGAGTGATT	600
QY	601	GACACATCAAAACCAACCTGCACACAGAGAATCATCATGTGGCTTGGAAATCCAAGA	660
Db	601	GACACATCAAAACCAACCTGCACACAGAGAATCATCATGTGGCTTGGAAATCCAAGA	660
QY	661	GATCTTCCTCTGATGGGCGCTCCATGCCATGCCCTTCGCCAGTTCATGTGGTGAAC	720
Db	661	GATCTTCCTCTGATGGGCGCTCCATGCCATGCCCTTCGCCAGTTCATGTGGTGAAC	720
QY	721	AGTGAAGTGTCTCTCCAGCTGTACAGAGATCGGAGACATGGGCGCTCGGTGCGCTTC	780
Db	721	AGTGAAGTGTCTCTCCAGCTGTACAGAGATCGGAGACATGGGCGCTCGGTGCGCTTC	780
QY	781	AACATGCGAGGTAGGCCCTGCTCAGTACATGATTTGGGCACATCACGGGCTGAGGCCA	840
Db	781	AACATGCGAGGTAGGCCCTGCTCAGTACATGATTTGGGCACATCACGGGCTGAGGCCA	840
QY	841	GGTGACTTTATACACATTTGGGAGATGACATATTTACCTGAAATCCATCGAGCCACTG	900
Db	841	GGTGACTTTATACACATTTGGGAGATGACATATTTACCTGAAATCCATCGAGCCACTG	900
QY	901	AAAAATTCAGCTTCAGCGAGAACCCAGACCTTTCCCAAGCTCAGGATTCCTCGAAAAGTT	960
Db	901	AAAAATTCAGCTTCAGCGAGAACCCAGACCTTTCCCAAGCTCAGGATTCCTCGAAAAGTT	960
QY	961	GAGAAAATGTATGACTTCAAAAGCTGAAGACTTTTCAGATTTGAAGGGTACAATCCGATCCA	1020
Db	961	GAGAAAATGTATGACTTCAAAAGCTGAAGACTTTTCAGATTTGAAGGGTACAATCCGATCCA	1020
QY	1021	ACTATTAAATGAAATGGCTGTTTAGGTGCTTTCAAAGGAGCTGGAAGATATTGTCA	1080
Db	1021	ACTATTAAATGAAATGGCTGTTTAGGTGCTTTCAAAGGAGCTGGAAGATATTGTCA	1080
QY	1081	GTTCTTAGGGGTGGGCTGGAGTCGAGGTAAAGTCTTTTGTCTCTAAAGAAAGG	1140
Db	1081	GTTCTTAGGGGTGGGCTGGAGTCGAGGTAAAGTCTTTTGTCTCTAAAGAAAGG	1140
QY	1141	AACTAGGTCAAAAATCTGTCGGTGACCTATCAGTTATTAAATTTTAAAGTATGTCGCAC	1200
Db	1141	AACTAGGTCAAAAATCTGTCGGTGACCTATCAGTTATTAAATTTTAAAGTATGTCGCAC	1200

QY	1201	GGCAAAATGTAACGTGCGCAGTCTCTTTCCATAATAAAGGCTTTGAGTTAACTCACTGAGG	1260
Db	1201	GGCAAAATGTAACGTGCGCAGTCTCTTTCCATAATAAAGGCTTTGAGTTAACTCACTGAGG	1260
QY	1261	GTATCTGCAAAATGCTGAGGTTAAGAACAAAGTGAGGAGAAATGAAATGATGTGCTCTTAG	1320
Db	1261	GTATCTGCAAAATGCTGAGGTTAAGAACAAAGTGAGGAGAAATGAAATGATGTGCTCTTAG	1320
QY	1321	CAAAAACATGTATGTGCAATTTCAATCCCAACGTACTATATAAAGAGGCTTGAGTAATTTCCAC	1380
Db	1321	CAAAAACATGTATGTGCAATTTCAATCCCAACGTACTATATAAAGAGGCTTGAGTAATTTCCAC	1380
QY	1381	AAGCTATATTTTGGAAATATTTTGTAGAAATATTTTAAAGAAATTTTCAACAAGCTATTTCCCTC	1440
Db	1381	AAGCTATATTTTGGAAATATTTTGTAGAAATATTTTAAAGAAATTTTCAACAAGCTATTTCCCTC	1440
QY	1441	CTGAGGAGAGCTGAGTAACACCATCGATCATGATGTAGAGTGTTGGTTATGAACTTTANAGT	1500
Db	1441	CTGAGGAGAGCTGAGTAACACCATCGATCATGATGTAGAGTGTTGGTTATGAACTTTANAGT	1500
QY	1501	TGTTTTATATGTCTCTATAATAAAGAGTGTTCTGCG	1536
Db	1501	TGTTTTATATGTCTCTATAATAAAGAGTGTTCTGCG	1536

RESULT 2

```

US-10-240-965-200
; Sequence 200, Application US/10240965
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: SHIFFMAN, Dov
; APPLICANT: SOMOGYI, Roland
; APPLICANT: LAWN, Richard M.
; APPLICANT: SEILHAMER, Jeffrey J.
; APPLICANT: PORTER, Gordon J.
; APPLICANT: MIKITA, Thomas
; APPLICANT: TAI, Julie
; TITLE OF INVENTION: GENES EXPRESSED IN FOAM CELL DIFFERENTIATION
; FILE REFERENCE: PA-0025 PCT
; CURRENT APPLICATION NUMBER: US/10/240,965
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: 60/195,106
; PRIOR FILING DATE: 2000-04-05
; NUMBER OF SEQ ID NOS: 276
; SOFTWARE: PERL Program
; SEQ ID NO 200
; LENGTH: 3298
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 247220.15
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 17, 1531, 1740-1765
; OTHER INFORMATION: a, t, c, g, or other
; US-10-240-965-200

```

Query Match	87.6%;	Score 1346;	DB 6;	Length 3298;
Best Local Similarity	99.9%;	Pred. No. 0;		
Matches 1446:	Conservative	0;	Mismatches 2;	Indels 0;
	Gaps	0;		

QY	49	GCCTGCTCTCGTCCGCGCGGCACATTGCCTGCTGCTGCTCCCGCGCGCGCCATG	108
Db	28	GCTTGCCTCTCGTCCGCGCGGCACATTGCCTGCTGCTGCTCCCGCGCGCGCCATG	87
QY	109	CCTGTGACCGGGTCGAGACTGCCGCGCGCGGCCCTTGTCCCCCGCCGCA CAGGAGCGGGAC	168
Db	88	CCTGTGACCGGGTCGAGACTGCCGCGCGCGGCCCTTGTCCCCCGCCGCA CAGGAGCGGGAC	147
QY	169	GCCGAGCGCGGTCGCGCGCACGGGAGCTGCAGTACTTGGGGGAGATCCACA CATCTTC	228

Db 361 CGACAGAGAAATCATATGTCGCTTGGAAATCCAAAGAGATCTTCTCTGATGGCGCTGCC 420

Qy 684 TCCATGCCATGCCCTCTGCGCAGTCTTATGTGTAACAGTGCAGTGTCTTGGCCAGCTGTA 743

Db 421 TCCATGCCATGCCCTCTGCGCAGTCTTATGTGTAACAGTGCAGTGTCTTGGCCAGCTGTA 480

Qy 744 CCAGAGATCGGAGACATGCGCCCTCGGTGTGCTTTCAACATCGCCAGCTTACGCCCTGCT 803

Db 481 CCAGAGATCGGAGACATGCGCCCTCGGTGTGCTTTCAACATCGCCAGCTTACGCCCTGCT 540

Qy 804 CAGTACATGATGTCGACATCATCGGCGCTTGAAGCCAGGTGACTTTATACACACTTTGGG 863

Db 541 CAGTACATGATGTCGACATCATCGGCGCTTGAAGCCAGGTGACTTTATACACACTTTGGG 600

Qy 864 AGATGCACATATTTACTCTGAAATCAATCATGAGCCACTGAAATTCAGTTCAGCGAGAACC 923

Db 601 AGATGCACATATTTACTCTGAAATCAATCATGAGCCACTGAAATTCAGTTCAGCGAGAACC 660

Qy 924 CAGACTTTCCCAAGCTCAGATTTCTTGGAAAGTTGAGAAAT 968

Db 661 CAGACTTTCCCAAGCTCAGATTTCTTGGAAAGTTGAGAAAT 705

RESULT 15

BQ057866

LOCUS

AGENCOURT 7047159 NIH_MGC_99 Homo sapiens cdna clone IMAGE:5813622

DEFINITION

5', mRNA sequence.

ACCESSION

BQ057866.1 GI:19817206

VERSION

EST.

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 1071)

AUTHORS

NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Lou Staudt
cdna Library Preparation: Rubin Laboratory
cdna Library Arrayed by: The I.M.A.G.E. Consortium (LILNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LILNL at:
<http://image.llnl.gov>
plate: LLCW2065 row: p column: 07
High quality sequence stop: 647.

FEATURES

source

1..1071

/organism="Homo sapiens"

/db_xref="taxon.9606"

/clone="IMAGE:5813622"

/clone_lib="NIH MGC 99"

/tissue_type="lymphoma, cell line"

/lab_host="DH10B (phage-resistant)"

/note="Organ: lymph, Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cdna made by oligo-dt priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: size GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

BASE COUNT

253 a 300 c 291 g 226 t 1 others

ORIGIN

Query Match 45.3%; Score 696; DB 14; Length 1071;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 696; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 TCCGCGCGCCACTTCCGCTGCTCCCTCCCGCCCGCCGCGCCATGCTGTGCGCCG 119

Db 1 TCCGCGCGCGCCACTTCCGCTGCTCCCTCCCGCCCGCCGCGCCATGCTGTGCGCCG 60

Qy 120 CTCGAGAGTTCGCGCGCGCGCCCTTTCGCGCGCGCGCGCGCAAGAGCGGAGCGCGAGCGCG 179

Db 61 CTCGAGAGTTCGCGCGCGCGCCCTTTCGCGCGCGCGCGCGCAAGAGCGGAGCGCGAGCGCG 120

Qy 180 TCCGCGCGCAAGCGGAGGTGAGTACCTGCGGCGAGATCCACACATCTCTCGCTGCGGCGGT 239

Db 121 TCCGCGCGCAAGCGGAGGTGAGTACCTGCGGCGAGATCCACACATCTCTCGCTGCGGCGGT 180

Qy 240 CAGGAAGAGCGAGCGAGCGTA 299

Db 181 CAGGAAGAGCGAGCGAGCGTA 240

Qy 300 CAGCCTGAGAGATGAATTCCTCTGCTGACAAACCAAGCTGTCTTCTGGAAGGCTGTCTTT 359

Db 241 CAGCCTGAGAGATGAATTCCTCTGCTGACAAACCAAGCTGTCTTCTGGAAGGCTGTCTTT 300

Qy 360 GGAGGAGTTGCTGTGTTTATCAAGGCGATCCAAATGCTTAAAGAGCTGTCTTCAAGGG 419

Db 301 GGAGGAGTTGCTGTGTTTATCAAGGCGATCCAAATGCTTAAAGAGCTGTCTTCAAGGG 360

Qy 420 AGTCAAAAATCTGGGATGCCAATGGATCCCGAGACTTTTGGACAGCGCTGGATCTCCAC 479

Db 361 AGTCAAAAATCTGGGATGCCAATGGATCCCGAGACTTTTGGACAGCGCTGGATCTCCAC 420

Qy 480 CAGAGAAGAGGCGGACTTGGGCGCGAGTTTATGCTTCCAGTCCAGTGGGCGATTTGGGGCAGA 539

Db 421 CAGAGAAGAGGCGGACTTGGGCGCGAGTTTATGCTTCCAGTGGGCGATTTGGGGCAGA 480

Qy 540 ATACAGAGATATGGATCAGATTATTCAGGACAGGCGAGTTGACCAACTGCAAAAGAGTGAT 599

Db 481 ATACAGAGATATGGATCAGATTATTCAGGACAGGCGAGTTGACCAACTGCAAAAGAGTGAT 540

Qy 600 TGACACCATCAAAACCAACCCCTGACGACAGAGAGATCATCATGTGCGCTTGGAAATCCAG 659

Db 541 TGACACCATCAAAACCAACCCCTGACGACAGAGAGATCATCATGTGCGCTTGGAAATCCAG 600

Qy 660 AGATCTTCTCTGATGGCGCTGCCCTCCATGCCATGCCCTCTGCCAGTTCTATGTGTGAA 719

Db 601 AGATCTTCTCTGATGGCGCTGCCCTCCATGCCATGCCCTCTGCCAGTTCTATGTGTGAA 660

Qy 720 CAGTGAAGCTCTCTGCGCGAGCTGTACAGAGATCGGG 755

Db 661 CAGTGAAGCTCTCTGCGCGAGCTGTACAGAGATCGGG 696

Search completed: November 29, 2002, 03:27:57

Job time : 2025.68 secs

```
RESULT 14
LOCUS      BG575942
DEFINITION 602598373F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:4706971 5',
            mRNA sequence.
ACCESSION  BG575942
VERSION     BG575942.1 GI:13583595
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Bkaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 888)
            NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS     National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE       Unpublished (1999)
JOURNAL
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTF
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10576 row: a column: 20
            High quality sequence stop: 867.

FEATURES             source
                     1..888
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4706971"
                     /clone_lib="NIH_MGC_87"
                     /tissue_type="mammary adenocarcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: breast; Vector: pCMV-SPORT6; Site 1: NotI;
                     Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.383 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC library."

BASE COUNT  245 a 198 c 232 g 213 t
ORIGIN
Query Match 45.9%; Score 705; DB 12; Length 888;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 705; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 264 CGGCACCTCTGCGTATTTCGGCATCGAGGCGCGCTACAGCCTGAGAGATGAATTCCTCT 323
DB 1 CGGCACCTCTGCGTATTTCGGCATCGAGGCGCGCTACAGCCTGAGAGATGAATTCCTCT 60
QY 324 GCTGACACCAACACGCTGTCTTCTGGAAGGGTGTCTTGGAGGAGTGTCTGTGTTTATCAA 383
DB 61 GCTGACACCAACACGCTGTCTTCTGGAAGGGTGTCTTGGAGGAGTGTCTGTGTTTATCAA 120
QY 384 GGGATCCACAAATGCTAAAGAGCTCTCTTCCAGGAGGTGAAATCTGGATGCCAATGG 443
DB 121 GGGATCCACAAATGCTAAAGAGCTCTCTTCCAGGAGGTGAAATCTGGATGCCAATGG 180
QY 444 ATCCCGAGACTTTTGGACAGCTGGGATCTCTCCACGAGAGAGAGGGGCTTGGGCCC 503
DB 181 ATCCCGAGACTTTTGGACAGCTGGGATCTCTCCACGAGAGAGAGGGGCTTGGGCCC 240
QY 504 AGTTTATGGCTTCCAGTGAGGCAATTTGGGCGAGATACAGAGATATGGAATCAGATTA 563
DB 241 AGTTTATGGCTTCCAGTGAGGCAATTTGGGCGAGATACAGAGATATGGAATCAGATTA 300
QY 564 TTCAGGACAGGGAGTTGACCACTGCAAGAGTGTATTGACACCATCAAAACCAACCTTGA 623
DB 301 TTCAGGACAGGGAGTTGACCACTGCAAGAGTGTATTGACACCATCAAAACCAACCTTGA 360
QY 624 CGACAGAGAGATCATCATCTGCGCTTGGAAATCCAGAGATCTTCTCTGATGGCGCTGCC 683
```

```
BASE COUNT  221 a 288 c 271 g 191 t 1 others
ORIGIN
Query Match 46.4%; Score 712; DB 14; Length 972;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 762; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 60 TCCGCGCGGCGGACCTTCGCTCCGTCCTCCCGCGCGCGCGCATGCTGTGGCCGG 119
DB 1 TCCGCGCGGCGGACCTTCGCTCCGTCCTCCCGCGCGCGCGCATGCTGTGGCCGG 60
QY 120 CTGCGAGCTGCGCGCGCGCGCTTCGCTCCCGCGCGCGCGCGCGCGCGCGCG 179
DB 61 CTGCGAGCTGCGCGCGCGCGCGCTTCGCTCCCGCGCGCGCGCGCGCGCGCG 120
QY 180 TCCGCGCGGCGGAGCTGCAGTACCTCTGGGCGAGATCCAAACATCTCTCGCTGCGGCGT 239
DB 121 TCCGCGCGGCGGAGCTGCAGTACCTCTGGGCGAGATCCAAACATCTCTCGCTGCGGCGT 180
QY 240 CAGGAAGAGACGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGAC 299
DB 181 CAGGAAGAGACGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGAC 240
QY 300 CAGCTGAGAGATGAATTCCTCTGCTGACACCAACGCTGTCTGGAAGGGTGTCTT 359
DB 241 CAGCTGAGAGATGAATTCCTCTGCTGACACCAACGCTGTCTGGAAGGGTGTCTT 300
QY 360 GGAGGAGTGTCTGTGTTTATCAAGGATCCAAATGCTTAAGAGCTGTCTTCCAAAGG 419
DB 301 GGAGGAGTGTCTGTGTTTATCAAGGATCCAAATGCTTAAGAGCTGTCTTCCAAAGG 360
QY 420 AGTCAAAATCTGGAGTGCCTATGATCCGAGACTTTTGGACAGCTGGCATCTCCAC 479
DB 361 AGTCAAAATCTGGAGTGCCTATGATCCGAGACTTTTGGACAGCTGGCATCTCCAC 420
QY 480 CAGAGAAGAGGAGACTTGGGCGCCAGTATTATGCTTCCAGTGGAGGCAATTTGGGGCAGA 539
DB 421 CAGAGAAGAGGAGACTTGGGCGCCAGTATTATGCTTCCAGTGGAGGCAATTTGGGGCAGA 480
QY 540 ATACAGAGATATGGAATCAGATTAATTCAGGACAGGGAGTTGACCACTGCAAGAGTGAT 599
DB 481 ATACAGAGATATGGAATCAGATTAATTCAGGACAGGGAGTTGACCACTGCAAGAGTGAT 540
QY 600 TGACACCATCAAAACCAACCCCTGACGACAGAGAAATCATCATGTGGCTTGGATCCAG 659
DB 541 TGACACCATCAAAACCAACCCCTGACGACAGAGAAATCATCATGTGGCTTGGATCCAG 600
QY 660 AGATCTTCTCTGATGGCGTGCCTCCATGCCATGCCCTCTGCGCATGCTATGTGGTAA 719
DB 601 AGATCTTCTCTGATGGCGTGCCTCCATGCCATGCCCTCTGCGCATGCTATGTGGTAA 660
QY 720 CAGTGAGCTGTCTGCGAGCTGTACAGAGATCGGAGACATGGGCTCGGCTGTCCTTT 779
DB 661 CAGTGAGCTGTCTGCGAGCTGTACAGAGATCGGAGACATGGGCTCGGCTGTCCTTT 720
QY 780 CAAATCGCAGGCTAGCCTCTGCTCAGTACATGATTTGGCGAC 822
DB 721 CAAATCGCAGGCTAGCCTCTGCTCAGTACATGATTTGGCGAC 763
```

[illegible]

	646	GCTTGGAAATCAAGAGATTCTTCCTCTGATGGCGGTGCCTCCATGCCATGCCCTCTGCCAG	705
Qy			
Ddb	481	GCTTGGAAATCCNAGAGATCTTCCTCTGATGGCGGTGCCTCCATGCCATGCCCTCTGCCAG	540
Qy	706	TTCATATGTGGTGAAACAGTAGAGCTGTCTCTGCCAGCTGTATACCAGAGATCGGAGACATGGGC	765
Ddb	541	TTCTATGTGGTGAAACAGTAGAGCTGTCTCTGCCAGCTGTATACCAGAGATCGGAGACATGGGC	600
Qy	766	CTCGGTGTGCCTTTCAAATCGCCAGCTACGCCCTGCTCAGTACATGATTGGCACATC	825
Ddb	601	CTCGGTGTGCCTTTCAAATCGCCAGCTACGCCCTGCTCAGTACATGATTGGCACATC	660
Qy	826	ACGGGCTGAAGCCAGCTGACTTTATACACACTTTGGGAGATGCACATATTTACCTGAAT	885
Ddb	661	ACGGGCTGAAGCCAGCTGACTTTATACACACTTTGGGAGATGCACATATTTACCTGAAT	720
Qy	886	CACATCGAGCCACTGAATAATTCAGCTTTCAGCGAGAACCCAGACCTTT	932
Ddb	721	CACATCGAGCCACTGAATAATTCAGCTTTCAGCGAGAACCCAGACCTTT	767
<hr/>			
RESULT 10			
BM904778			
LOCUS			
DEFINITION			
5', mRNA sequence.			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
NIH-MGC http://mgc.nci.nih.gov/ National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgabbs-rc@mail.nih.gov Tissue procurement: ATCC/DCTP cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAML2279 row: c column: 24 High quality sequence stop: 673.			
<hr/>			
FEATURES			
source			
1..994			
/organism="Homo sapiens"			
/db xref="taxon:9606"			
/clone="IMAGE:5557583"			
/clone_lib="NIH MGC 72"			
/tissue_type="melanotic melanoma"			
/lab_host="DH10B (phage-resistant)"			
/note="Organ: skin; Vector: pCMV-Sport6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr. Average insert size 2 kb. Library constructed by Life Technologies."			
<hr/>			
BASE COUNT			
ORIGIN			
Query Match			
Best Local Similarity 48.5%; Score 745; DB 14; Length 994;			
Matches 745; Conservative 0; Mismatches 0; Indels 0; Gaps			
<hr/>			
Qy	45	CTTTGGCTGCTCTGTCGCGCGGCACATTCGCTGCTCGCTCCGTCCTCCCGCGCGCGCGCG	104
Ddb	1	CTTTGGCTGCTCTGTCGTCGCGCGGCACATTCGCTGCTCTCGCTCCGTCCTCCCGCGCGCGCGCG	60
Qy	105	CATGCGCTTGGCGCGGCTCGGAGCTGCGCGCGCGGCCCTTGGCCCCCGCGCGCACAGGAGCG	164

Db	254	CAAGCGCGCTACAGCTGAGAGATGAATTCCTCTGCTGACCAACAAAGCTGTCTTCGCG	313
QY	349	AAGGGTGTTTGAGGAGTTCCTGTGTTTATCAAGAGATCCCAAAATGCTTAAAGAGCTG	408
Db	314	AAAGGTGTTTTGGAGAGTTCGTGTGTATCAAGGATCCAAATGCTTAAAGAGCTG	373
QY	409	TCCTCCAAAGGGAGTGAATCTGGGATGCCAATGATCCCGAGACTTTTGGACAGAGCTG	468
Db	374	TCCTCCAAAGGGAGTGAATCTGGGATGCCAATGATCCCGAGACTTTTGGACAGAGCTG	433
QY	469	GGATTTCCACAGACAGAAAGGGGACTTGGGCGCAGTTTATGCGTTCCAGTGGAGGCAT	528
Db	434	GGATTTCTCCACAGACAGAGGGGACTTGGGCGCAGTTTATGCGTTCCAGTGGAGGCAT	493
QY	529	TTTGGGGCAGAAATACAGACATATGGAATCAAGATTTATCAGACAGGGAGTTGACCACTG	588
Db	494	TTTGGGGCAGAAATACAGACATATGGAATCAAGATTTATCAGACAGGGAGTTGACCACTG	553
QY	589	CAAGAGTGATTTGACACCAATCAAAACCAAGCTGTGAGGCAGAGAAATCATCATGTGGCT	648
Db	554	CAAGAGTGATTTGACACCAATCAAAACCAAGCTGTGAGGCAGAGAAATCATCATGTGGCT	613
QY	649	TGGATCCACAGAGACTTTCCTCTGATGGCGTGCCTCCATGCCATGCGCTCTGCCAGTTC	708
Db	614	TGGAATCCACAGAGACTTTCCTCTGATGGCGTGCCTCCATGCCATGCGCTCTGCCAGTTC	673
QY	709	TATGTGTTGAACAGTCAGCTGTCCTGCGACGTGTACCAAGATCGGAGACATGGGGCTC	768
Db	674	TATGTGTTGAACAGTCAGCTGTCCTGCGACGTGTACCAAGATCGGAGACATGGGGCTC	733
QY	769	GGTGTGCGCTTTCAACATCCCGAGCTACGCCCTCTCTCAGTACATGATTCGGCAATCAAG	828
Db	734	GGTGTGCGCTTTCAACATCCCGAGCTACGCCCTCTCTCAGTACATGATTCGGCAATCAAG	793
QY	829	GGCCTGAAGCCAGG	842
Db	794	GGCCTGAAGCCAGG	807

RESULT 8	
BM473957	
LOCUS	1035 bp mRNA linear EST 05-FEB-2002
DEFINITION	AGENCOURT_6467897 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:582212
ACCESSION	BM473957
VERSION	BM473957.1 GI:18522999
KEYWORDS	5', mRNA sequence.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 1035)
TITLE	NIH-MGC http://mgs.nci.nih.gov/ .
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT	Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgapbs- reemail.nih.gov Tissue Procurement: ATCC cDNA library preparation: Life Technologies, Inc. cDNA library arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ LLNL at: http://image.llnl.gov Plate: LLAM12343 row: f column: 05 High quality sequence stop: 687. Location/Qualifiers 1..1035 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:582212" /clone_lib="NIH MGC 67"
FEATURES	
source	

/tissue type="placenta"
/notes="Vector: pCMVSPORT 6; site 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"
BASE COUNT 220 a 263 c 254 g 193 t 2 others
ORIGIN

Query Match 52.0%; Score 799; DB 9; Length 932;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 899; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 66 CAGGCACTTGCCTGCTCGTCCCGCCCGCCGCGCCATGCTGTGGCGGCTCGGA 125
Db 1 CCGCGCACTTGCCTGCTCGTCCCGCCCGCCGCGCCATGCTGTGGCGGCTCGGA 60
QY 126 GCTGCGCGCGCGCCCTTGGCCCGCCCGCCGCGAGCGGAGCGCGCTCCGCC 185
Db 61 GCTGCGCGCGCGCCCTTGGCCCGCCCGCCGCGAGCGGAGCGCGCTCCGCC 120
QY 186 GCACGGGAGCTGCAGTACCTTGGGCGAGATCCACACATCTCTGCGCGTACAGAA 245
Db 121 GCACGGGAGCTGCAGTACCTTGGGCGAGATCCACACATCTCTGCGCGTACAGAA 180
QY 246 GGAACGACGACCGGCGACCGGCGACCTGTGCGTATTCGGATGACGCGCGCTACAGCT 305
Db 181 GGAACGACGACCGGCGACCGGCGACCTGTGCGTATTCGGATGACGCGCGCTACAGCT 240
QY 306 GAGAGATGAATCTCTGCTGACACCAACGCTGTGCTTGGAGGCTTTTGGAGGA 365
Db 241 GAGAGATGAATCTCTGCTGACACCAACGCTGTGCTTGGAGGCTTTTGGAGGA 300
QY 366 GTTGTGTGTTTATCAAGGATCCAAATGTAAGAGCTGTCTTCCAAGGAGTGAA 425
Db 301 GTTGTGTGTTTATCAAGGATCCAAATGTAAGAGCTGTCTTCCAAGGAGTGAA 360
QY 426 AATCTGGGATCCAAATGATCCGAGACTTTTGGACAGCTGGGATTTCCACCGAGA 485
Db 361 AATCTGGGATCCAAATGATCCGAGACTTTTGGACAGCTGGGATTTCCACCGAGA 420
QY 486 AGAAGGGAGCTTGGGCCCGAGTTATGGCTTCCAGTGAGGAGATTTGGGCGAGATACAG 545
Db 421 AGAAGGGAGCTTGGGCCCGAGTTATGGCTTCCAGTGAGGAGATTTGGGCGAGATACAG 480
QY 546 AGATATGAATCAGATTTACAGACAGGAGTTGACCACTGACCAAGAGTGATTCACAC 605
Db 481 AGATATGAATCAGATTTACAGACAGGAGTTGACCACTGACCAAGAGTGATTCACAC 540
QY 606 CATCAAAACCAACCTGACGACAGAGATCATGTGCGCTTGGATCCAGAGATCT 665
Db 541 CATCAAAACCAACCTGACGACAGAGATCATGTGCGCTTGGATCCAGAGATCT 600
QY 666 TCCTGTGATGGCGTGTCTCCATGCCATGCCCTCTGCCAGTTCTATGTGTGACAGTGA 725
Db 601 TCCTGTGATGGCGTGTCTCCATGCCATGCCCTCTGCCAGTTCTATGTGTGACAGTGA 660
QY 726 GCTGTCTCCAGCTGTATACAGAGATCGGAGACATGCGGCTCTGCTGTGCTTCAACAT 785
Db 661 GCTGTCTCCAGCTGTATACAGAGATCGGAGACATGCGGCTCTGCTGTGCTTCAACAT 720
QY 786 CGCAGCTACGCGCTGTCTACGATACATGATGCGACATCATCGGCGCTGAGGCGAGTGA 845
Db 721 CGCAGCTACGCGCTGTCTACGATACATGATGCGACATCATCGGCGCTGAGGCGAGTGA 780
QY 846 CTTTATACACATTTGGGAGATGACATATTTACTGATATCATCATGAGGCGACTGAAAT 905
Db 781 CTTTATACACATTTGGGAGATGACATATTTACTGATATCATCATGAGGCGACTGAAAT 840

QY 906 TCAGTTTCAGCGACGACCCAGACCTTTCCCAAGCTCAGGATTTCTTCGAAAAGTTGAGAA 965
Db 841 TCAGTTTCAGCGACGACCCAGACCTTTCCCAAGCTCAGGATTTCTTCGAAAAGTTGAGAA 900
QY 966 A 966
Db 901 A 901

RESULT 7
BQ420981
LOCUS BQ420981 889 bp mRNA linear EST 23-MAY-2002
DEFINITION AGSCOURT_7761945 NIH_MGC_70 Homo sapiens cDNA clone IMAGE:6018140
5', mRNA sequence.
ACCESSION BQ420981
VERSION BQ420981.1 GI:21116296
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 889)
NIH-MGC http://mgi.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13218 row: i column: 21
High quality sequence stop: 715.

FEATURES
Location/Qualifiers
1..889
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6018140"
/clone_lib="NIH MGC 70"
/tissue_type="epithelioid carcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: pancreas; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.1 kb. Library constructed by Life Technologies."
BASE COUNT 195 a 265 c 248 g 181 t
ORIGIN

Query Match 51.7%; Score 794; DB 14; Length 889;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 794; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 GCCTGTGCTCGCTCCGCGCGACCTTCCTGCTCCCTCCCGCCCGCGCGCGCATG 108
Db 14 GCCTGTGCTCGCTCCGCGCGACCTTCCTGCTCCCTCCCGCCCGCGCGCATG 73
QY 109 CCTGTGCGCGGCTCGAGCTGCGCGCGCGCCCTTCGCCCGCCGCGACAGGAGCGGGAC 168
Db 74 CCTGTGCGCGGCTCGAGCTGCGCGCGCGCCCTTCGCCCGCCGCGACAGGAGCGGGAC 133
QY 169 GCCGAGCGCGCTCCGCGCGACGCGGAGCTGACGACTCTGGGCGAGATCCAAACATCTC 228
Db 134 GCCGAGCGCGCTCCGCGCGACGCGGAGCTGACGACTCTGGGCGAGATCCAAACATCTC 193
QY 229 CGCTGCGCGCTCAGGAGAGGACGACCGACCGGCGACCGGCGACCTGTGCGTATTCGCGATG 288
Db 194 CGCTGCGCGCTCAGGAGAGGACGACCGACCGGCGACCGGCGACCTGTGCGTATTCGCGATG 253
QY 289 CAGGCGCGCTACAGCGCTGAGAGATGAATTCCTCTGCTGACCAACCAACGCTGTTCTGG 348

```

Db 661 CTGCGAGCTGTACAGAGATCGGAGACATGGGCTCGGTGCTTTCAACATCGCA 720
|
|
|
Qy 791 GCTACGCCCTGCTCAGTACATGATGTCGACATACAGGCGCTGAAGCCAGTGACTTTA 850
|
|
|
Db 721 GCTACGCCCTGCTCAGTACATGATGTCGACATACAGGCGCTGAAGCCAGTGACTTTA 780
|
|
|
Qy 851 TACACACTTTGGGAGATCAGATATTTACCTGATCATCATGAGCCACTGAAATTCAGC 910
|
|
|
Db 781 TACACACTTTGGGAGATCAGATATTTACCTGATCATCATGAGCCACTGAAATTCAGC 840
|
|
|
Qy 911 TTTACGCGAGAACCC 924
|
|
|
Db 841 TTTACGCGAGAACCC 854
|
|
|
RESULT 5
BM478045 1045 bp mRNA linear EST 05-FEB-2002
LOCUS AGENCOURT_6482920 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:5552680
5', mRNA sequence.
ACCESSION BM478045
VERSION BM478045.1 GI:18527087
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1045)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Lou Staudt
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12268 row: 9 column: 17
High quality sequence stop: 687.
FEATURES
Location/Qualifiers
1..1045
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5552680"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="organ: lymph; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.967 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC library."
BASE COUNT 242 a 297 c 286 g 220 t
ORIGIN
Query Match 53.7%; Score 825; DB 13; Length 1045;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 825; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 33 CCCGCGCGGCACATGGCGCTGCTCCGTCGCGCGCCACCTTCGCTGCTCGGTCGCC 92
Db 1 CCCGCGCGGCACATGGCGCTGCTCCGTCGCGCGCCACCTTCGCTGCTCGGTCGCC 60
Qy 93 CGCCCGCGCGGCACATGGCTGTGCGCGGCTGCGAGCTGCGCGCGCGCTTCGCCCGCC 152
Db 61 CGCCCGCGCGGCACATGGCTGTGCGCGGCTGCGAGCTGCGCGCGCGCTTCGCCCGCC 120
Qy 153 CGCACAGAGCGGAGCGCGCGCTGCGCGGCTGCGAGCTGCGCGAGCTGCACTACTGGGCA 212
|
|
|

```

```

Db 121 CGCACAGGAGCGGAGCGCGAGCCGCGTCCGCGCAGCGGGAGCTGCAGTACCTGGGGCA 180
|
|
|
Qy 213 GATCCAAACATCTCTCCGCTCGGCGGTGAGGAGGACGACCGCAGGGGACCGGACACCT 272
|
|
|
Db 181 GATCCAAACATCTCTCCGCTCGGCGGTGAGGAGGACGACCGCAGGGGACCGGACACCT 240
|
|
|
Qy 273 GTCGGTATTCGGCATGCAAGCGCGCTACAGCTGAGAGATGAATTCCTCTGCTGACAAAC 332
|
|
|
Db 241 GTCGGTATTCGGCATGCAAGCGCGCTACAGCTGAGAGATGAATTCCTCTGCTGACAAAC 300
|
|
|
Qy 333 CAAACGCTGCTTCTGGAAGGCTGTTTGGAGAGTGTCTGTGTGTTTATCAAGGGATCCAC 392
|
|
|
Db 301 CAAACGCTGCTTCTGGAAGGCTGTTTGGAGAGTGTCTGTGTGTTTATCAAGGGATCCAC 360
|
|
|
Qy 393 AAATGCTAAAGAGCTGTCTTCCAGGGAGTGAAATCTGGATGCGCAATGGATCCCGAGA 452
|
|
|
Db 361 AAATGCTAAAGAGCTGTCTTCCAGGGAGTGAAATCTGGATGCGCAATGGATCCCGAGA 420
|
|
|
Qy 453 CTTTTCGAGCAGCTGGGATTTCTCCACAGAGAGAGGGGACTTGGGCCCAGTTTATGG 512
|
|
|
Db 421 CTTTTCGAGCAGCTGGGATTTCTCCACAGAGAGAGGGGACTTGGGCCCAGTTTATGG 480
|
|
|
Qy 513 CTTCCAGTGGAGGCAATTTTGGGGCAGATATCAGAGATATGGAATCAGATTATTCAGGACA 572
|
|
|
Db 481 CTTCCAGTGGAGGCAATTTTGGGGCAGATATCAGAGATATGGAATCAGATTATTCAGGACA 540
|
|
|
Qy 573 GGGAGTTGACCAACTGCAAGAGTGATTTGACCAATCAAAACCAACCTTCGACGACAGAG 632
|
|
|
Db 541 GGGAGTTGACCAACTGCAAGAGTGATTTGACCAATCAAAACCAACCTTCGACGACAGAG 600
|
|
|
Qy 633 AATCATCATGTGCGCTTGAATTCGAAGAGATCTTCCTGTGATGGGCGCTGCCCTCCATGCCA 692
|
|
|
Db 601 AATCATCATGTGCGCTTGAATTCGAAGAGATCTTCCTGTGATGGGCGCTGCCCTCCATGCCA 660
|
|
|
Qy 693 TGGCCTCTGCCAGTTCTATGTGTGTGAACAGTGAGCTGTCTGCGAGCTGTACCCAGAGATC 752
|
|
|
Db 661 TGGCCTCTGCCAGTTCTATGTGTGTGAACAGTGAGCTGTCTGCGAGCTGTACCCAGAGATC 720
|
|
|
Qy 753 GGGAGATATGGCGCTGGGTGCTTTCACATCGCAGCTGCTGCGAGCTGCTCAGCTACAT 812
|
|
|
Db 721 GGGAGATATGGCGCTGGGTGCTTTCACATCGCAGCTGCTGCGAGCTGCTCAGCTACAT 780
|
|
|
Qy 813 GATTGCGCATCATCAGCGGCTGAGCCAGGTGACTTTATATACAC 857
|
|
|
Db 781 GATTGCGCATCATCAGCGGCTGAGCCAGGTGACTTTATATACAC 825
|
|
|
RESULT 6
AL551817 932 bp mRNA linear EST 16-FEB-2001
LOCUS AL551817 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0D1061YF09 5
prime, mRNA sequence.
ACCESSION AL551817
VERSION AL551817.1 GI:12890127
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 932)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
Location/Qualifiers
1..932
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D1061YF09"
/clone_lib="LTI_NFL006_PL2"
source

```

74 TTGGCTGCTCTGCTCCCGCCCGCCGCGCCGCGCCATGCTGTGCGCGCTCGAGCTCGCG 133
 Db TTGGCTGCTCTGCTCCCGCCCGCCGCGCCGCGCCATGCTGTGCGCGCTCGAGCTCGCG 126

134 GCCGGCCCTTGGCCCGCCCGCCGCGCCGCGCCGCGCCGCGCCGCGCCGCGCCGCGCG 193
 Db GCCGGCCCTTGGCCCGCCCGCCGCGCCGCGCCGCGCCGCGCCGCGCCGCGCGCGCG 186

194 AGCTGCACTGCTGCG 253
 Db AGCTGCACTGCTGCG 246

254 GCACGGCG 313
 Db GCACGGCG 306

314 AATTCCTCTGCTGACCAACCAACCGTGTGTTCTGGAAGCGTGTGTTGAGGAGTGTCTGT 373
 Db AATTCCTCTGCTGACCAACCAACCGTGTGTTCTGGAAGCGTGTGTTGAGGAGTGTCTGT 366

374 GCTTATCAAGGATCCACAAATGCTAAAGAGCTGTCTTCAAGGAGTGAATAATCTGGG 433
 Db GCTTATCAAGGATCCACAAATGCTAAAGAGCTGTCTTCAAGGAGTGAATAATCTGGG 426

434 ATGCCAATGATCCCGGAGACTTTTTCGACAGCTCGGATCTTCCACGAGAGAGAGGG 493
 Db ATGCCAATGATCCCGGAGACTTTTTCGACAGCTCGGATCTTCCACGAGAGAGAGGG 486

494 ACTTGGGCGCCAGTTTATGGCTTCCAGTGGAGGCAATTTTGGGCGAGATACAGAGATATGG 553
 Db ACTTGGGCGCCAGTTTATGGCTTCCAGTGGAGGCAATTTTGGGCGAGATACAGAGATATGG 546

554 AATCAGATTAATCAGGACAGGAGTTGACCAATCTGCAAGAGTGAATTGACACCATCAAAA 613
 Db AATCAGATTAATCAGGACAGGAGTTGACCAATCTGCAAGAGTGAATTGACACCATCAAAA 606

614 CCAACCTGAGCAGAGAGATCATCATGCGCTGGAATCCAGAGATCTTCTCTCTGA 673
 Db CCAACCTGAGCAGAGAGATCATCATGCGCTGGAATCCAGAGATCTTCTCTCTGA 666

674 TGGCGCTGCTCCATGCGCTCTGCGCTCTGCGCTCTGCGCTCTGCGCTCTGCGCTCTGCT 733
 Db TGGCGCTGCTCCATGCGCTCTGCGCTCTGCGCTCTGCGCTCTGCGCTCTGCGCTCTGCT 726

734 GCGAGCTTACAGAGATCGGAGACATGCGCTGCGCTGCGCTGCGCTTCAACATCGCCAGCT 793
 Db GCGAGCTTACAGAGATCGGAGACATGCGCTGCGCTGCGCTGCGCTTCAACATCGCCAGCT 786

794 AGCCCTGCTCACTGATCATGATGCGCAGATCATGCGCTGCGCTGCGCTGCGCTGCGCT 853
 Db AGCCCTGCTCACTGATCATGATGCGCAGATCATGCGCTGCGCTGCGCTGCGCTGCGCT 846

854 ACACCTTTGGGAGATGCACATATTTACTGAAATCAATCGAGGCGCACTGA 901
 Db ACACCTTTGGGAGATGCACATATTTACTGAAATCAATCGAGGCGCACTGA 894

RESULT 4
 AL556369
 LOCUS
 DEFINITION AL556369 LTI_NFL006 PL2 Homo sapiens cDNA clone CS0DK004YL08 5
 prime, mRNA sequence.
 ACCESSION AL556369
 VERSION AL556369.1 GI:12898981
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 854)
 Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)

COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
 Location/Qualifiers
 1. .854
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CS0DK004YL08"
 /clone_lib="LTI_NFL006_PL2"
 /issue_type="placenta"
 /note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA
 was primed with a NotI-oligo(dT) primer. Five prime end
 enriched, double-stranded cDNA was digested with Not I and
 cloned into the Not I and Eco RV sites of the pCMVSPORT 6
 vector. Library was normalized. Library was constructed by
 Life Technologies. Contact : Feng Liang Life Technologies,
 a division of Invitrogen 9800 Medical Center Drive
 Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
 Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com"

BASE COUNT 194 a 247 c 239 g 174 t
 ORIGIN

Query Match 55.6%; Score 854; DB 9; Length 854;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 854; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 71 CACTTCGCTGCTCTGCTCCCGCCCGCCGCGCCGCGCCATGCTGTGCGCGCTCGAGCTGC 130
 Db 1 CACTTCGCTGCTCTGCTCCCGCCCGCCGCGCCGCGCCATGCTGTGCGCGCTCGAGCTGC 60

QY 131 CGCGCGCGCGCTTGGCCCGCCCGCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 190
 Db 61 CGCGCGCGCGCTTGGCCCGCCCGCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120

QY 191 GGGAGCTGCACTGCTGCG 250
 Db 121 GGGAGCTGCACTGCTGCG 180

QY 251 ACCGCAAGGCG 310
 Db 181 ACCGCAAGGCG 240

QY 311 ATGAATTCCTCTGCTGACCAACCAACCGTGTGTTCTGGAAGGCTGTTTGGAGGAGTGC 370
 Db 241 ATGAATTCCTCTGCTGACCAACCAACCGTGTGTTCTGGAAGGCTGTTTGGAGGAGTGC 300

QY 371 TGTGGTTTATCAAGGATCCACAAATGCTAAAGAGCTGTCTTCAAGGAGTGAATCT 430
 Db 301 TGTGGTTTATCAAGGATCCACAAATGCTAAAGAGCTGTCTTCAAGGAGTGAATCT 360

QY 431 GGGATGCCAATGATCCCGGAGCTTTTGGAGAGCTTTTGGAGAGCTTTTGGAGAGAGAG 490
 Db 361 GGGATGCCAATGATCCCGGAGCTTTTGGAGAGCTTTTGGAGAGCTTTTGGAGAGAGAG 420

QY 491 GGGACTTGGGCG 550
 Db 421 GGGACTTGGGCG 480

QY 551 TGGATCAGATTAATTCAGGACAGGAGTGAACCAATGCAAGAGTGAATGACCAATCA 610
 Db 481 TGGATCAGATTAATTCAGGACAGGAGTGAACCAATGCAAGAGTGAATGACCAATCA 540

QY 611 AAACCAACCTGACGACAGAGATCATGTCGCTTGGATCCAGAGATCTTCTCTC 670
 Db 541 AAACCAACCTGACGACAGAGATCATGTCGCTTGGATCCAGAGATCTTCTCTC 600

QY 671 TGATGGCGCTGCTTCCATCCCTGCTTCCGAGTTCCTATGCTGAGACAGTGAATCT 730
 Db 601 TGATGGCGCTGCTTCCATCCCTGCTTCCGAGTTCCTATGCTGAGACAGTGAATCT 660

QY 731 CTTGCCAGCTGTACAGAGATCGGAGAGATGCGGCTGCTGCTGCTTCAACATCGCCA 790

Shvchenko, Y., Metherby, K.D., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Jagspi, R., Lim, M., Maduro, Q.L., Mastello, C., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Snyder, B., Stanthrop, S., Thomas, P.J., Thompson, E.B., Touchman, J.W., Tsugeon, C., Vogt, J.L., Walker, M.A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILM at: <http://image.llnl.gov>
Series: IRM1 Plate: 4 Row: h Column: 10
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 10635001
This clone has the following problem: no 5' EST match.

FEATURES

Location/Qualifiers
1..1533

Source

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3141419"
/tissue_type="Ovary, adenocarcinoma"
/clone_lib="NIH MGC 9"
/lab_host="DH10B-R"
/note="Vector: pOT87"

BASE COUNT 424 a 354 c 382 g 373 t

ORIGIN

Query Match 84.0%; Score 1291; DB 11; Length 1533;
Best Local Similarity 99.8%; Pred. No. 0;

Matches 1441; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

53 GCTTCCTCCGCGCGGCGCACTTCGCTCCGCTCCCGCGCGCGCGCATGCTG 112
9 GCTTCCTCCGCGCGGCGCACTTCGCTCCGCTCCCGCGCGCGCGCATGCTG 68
113 TGGCCGCGCTGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCATGCTG 172
69 TGGCCGCGCTGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCATGCTG 128
173 AGCCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCATGCTG 232
129 AGCCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCATGCTG 188
233 GCGCGCTGAGGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCATGCTG 292
189 GCGCGCTGAGGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCATGCTG 248
293 GCGCGCTGAGGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCATGCTG 352
249 GCGCGCTGAGGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCATGCTG 308
353 GTGTTTGGAGAGGAGTGTGTTTATCAAGGATCAACAATGCTAAAGAGCTGCTT 412
309 GTGTTTGGAGAGGAGTGTGTTTATCAAGGATCAACAATGCTAAAGAGCTGCTT 368
413 CCAAGGAGTGAATATCGGATGCGCAATGATCCGAGATTCTTTTGAACGCTGGAT 472
369 CCAAGGAGTGAATATCGGATGCGCAATGATCCGAGATTCTTTTGAACGCTGGAT 428
473 TCTCCACCAAGAAAGAGGAGCTTGGGCGCATTTATGCTTCAGTGGAGCATTTTG 532
429 TCTCCACCAAGAAAGAGGAGCTTGGGCGCATTTATGCTTCAGTGGAGCATTTTG 488
533 GGGCGAATACAGATATGGAATCAATATTCAGGACAGGAGTTGACCAATCGGAA 592
489 GGGCGAATACAGATATGGAATCAATATTCAGGACAGGAGTTGACCAATCGGAA 548
593 GAGGATTTGACACATCAAAACCAACCTGACGACAGAGAAATCAATCAATGCTGTA 652
549 GAGGATTTGACACATCAAAACCAACCTGACGACAGAGAAATCAATCAATGCTGTA 608
653 ATCCAAAGATCTTCTGATGGCGTCCCTCCATGCGCAGTCCCTCTGAGTTCTATG 712
609 ATCCAAAGATCTTCTGATGGCGTCCCTCCATGCGCAGTCCCTCTGAGTTCTATG 668

713 TGTGAACAGTGAAGTGTCTGCGAGCTGTACAGAGATCGGAGACATGCGCTCGGTG 772
669 TGTGAACAGTGAAGTGTCTGCGAGCTGTACAGAGATCGGAGACATGCGCTCGGTG 728
773 TGCCTTCAACATCGCAGCTACGCGCTGCTACAGTACATGATTCGCGACATCGGCG 832
729 TGCCTTCAACATCGCAGCTACGCGCTGCTACAGTACATGATTCGCGACATCGGCG 788
833 TGAAGCCAGTGTATTAACACACTTGGGAGATGACATATTAACCTGAATCAGATCG 892
789 TGAAGCCAGTGTATTAACACACTTGGGAGATGACATATTAACCTGAATCAGATCG 848
893 AGCCATGAAATATGAGCTTACGAGAGAACCCAGACCTTCCCAAGCTCAGATCTTC 952
849 AGCCATGAAATATGAGCTTACGAGAGAACCCAGACCTTCCCAAGCTCAGATCTTC 908
953 GAAAGTTGAGAAATATGATGATCAATCAAGCTGAGACCTTCAATGAGAGGATCAATC 1012
909 GAAAGTTGAGAAATATGATGATCAATCAAGCTGAGACCTTCAATGAGAGGATCAATC 968
1013 CGCATCCACTATTAATGAATGAAATGCTGTTAGGAGCTTTCAAGGAGCTMGAAAGA 1072
969 CGCATCCACTATTAATGAATGAAATGCTGTTAGGAGCTTTCAAGGAGCTMGAAAGA 1028
1073 TATGTGAGTCTTTAGGAGGTTGGGCTGATGCGAGGTAAGTCTTTGCTCTTAA 1132
1029 TATGTGAGTCTTTAGGAGGTTGGGCTGATGCGAGGTAAGTCTTTGCTCTTAA 1088
1133 GAANAAGAACTGATGCTAAATATGCTGACCTTCAATGATTAATTTTAAGATG 1192
1089 GAANAAGAACTGATGCTAAATATGCTGACCTTCAATGATTAATTTTAAGATG 1148
1193 TTGCACTGCGCAATGATGATGCTGCAATCTTCCATTAATTAAGCTTTGATTAAT 1252
1149 TTGCACTGCGCAATGATGATGCTGCAATCTTCCATTAATTAAGCTTTGATTAAT 1208
1253 CACTGAGGATATGCAATGATGATGATTAAGAAAGTGAAGAGATGAATGATG 1312
1209 CACTGAGGATATGCAATGATGATGATTAAGAAAGTGAAGAGATGAATGATG 1268
1313 GCTTTAGCAAAAATCATGATGCTGATTTCAATCCAGTCTTAATTAAGAGTTGG 1372
1269 GCTTTAGCAAAAATCATGATGCTGATTTCAATCCAGTCTTAATTAAGAGTTGG 1328
1373 AATTTCACAGCTATTTTGAATATTTTGAATATTTTAAGATTTTCAAGCTATTC 1432
1329 AATTTCACAGCTATTTTGAATATTTTGAATATTTTGAATATTTTCAAGCTATTC 1388
1433 CCTCAATCTGAGGAGCTGATGACACATGATGATGATGATGATGATGATGATG 1492
1389 CCTCAATCTGAGGAGCTGATGACACATGATGATGATGATGATGATGATGATG 1448
1493 TTTA 1496
1449 TTTA 1452

RESULT 2
AL547612 1020 bp mRNA linear EST 16-FEB-2001
LOCUS AL547612 LTI.NFL006.PL2 Homo sapiens cDNA clone CS010087H21 5
DEFINITION prime, mRNA sequence.
ACCESSION AL547612
VERSION AL547612.1 GI:12881832
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1020)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: November 29, 2002, 00:16:50 ; Search time 2021.68 Seconds
(without alignments)
12304.774 Million cell updates/sec

Title: US-09-963-333-6
Perfect score: 1536
Sequence: 1 ggggggggggggggaccacttg.....ataataagaagtgtctgc 1536

Scoring table: OLIGO_NUC
Gapop_60.0 , Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues

Word size : 15
Total number of hits satisfying chosen parameters: 36158

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST.*
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estnu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_hic.*
9: gb_est1.*
10: gb_est2.*
11: gb_hic.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: gb_gss.*
18: em_gss_hum.*
19: em_gss_inv.*
20: em_gss_pln.*
21: em_gss_vrt.*
22: em_gss_fun.*
23: em_gss_man.*
24: em_gss_mus.*
25: em_gss_other.*
26: em_gss_pro.*
27: em_gss_rod.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1291	84.0	1533	11	BC018858 Homo sapi
2	985	64.1	1020	9	AL547612 AL547612
3	888	57.8	938	9	AL551990 AL551990
4	854	55.6	854	9	AL556369 AL556369
5	825	53.7	1045	13	BM478045 AGENCOURT
6	799	52.0	932	9	AL551817 AL551817

7	794	51.7	889	14	BQ420981
8	781	50.8	1035	13	BM473957
9	767	49.9	855	9	AU142370
10	745	48.5	994	14	BM904778
11	736	47.9	948	14	BQ068205
12	731	47.6	935	9	AL542409
13	712	46.4	972	14	BQ642223
14	705	45.9	888	12	BG575942
15	696	45.3	1071	14	BQ057866
16	695	45.2	903	13	BM459789
17	690	44.9	1089	9	AL575295
18	688	44.8	1033	14	BM928567
19	687	44.7	742	9	AL548721
20	682	44.4	722	12	BG574546
21	682	44.4	900	13	BM453660
22	675	43.9	1046	14	BM919886
23	669	43.6	1072	13	BM475928
24	657	42.8	809	9	AU117646
25	656	42.7	1072	14	BQ066271
26	653	42.5	904	14	BQ642214
27	651	42.4	841	14	BQ434233
28	650	42.3	1149	13	BM546045
29	646	42.1	907	14	BQ223405
30	643	41.9	888	13	BM413631
31	641	41.7	939	13	BM452851
32	638	41.5	1135	14	BQ062012
33	634	41.3	893	13	BM457044
34	630	41.0	987	13	BM472540
35	629	41.0	789	13	BI871083
36	623	40.6	904	14	BQ669537
37	623	40.6	970	14	BQ928426
38	617	40.2	741	9	AU118644
39	614	40.0	984	14	BQ061952
40	613	39.9	890	12	BG107998
41	609	39.6	891	14	BQ224163
42	602	39.2	825	9	AU117269
43	602	39.2	906	9	AL578646
44	595	38.7	738	9	AU130205
45	591	38.5	591	14	BM843120

ALIGNMENTS

BC018858 Homo sapiens, clone IMAGE:3141419, mRNA. linear HTC 07-DEC-2001
1533 bp

BC018858.1 GI:17402996

HTC.

HTC.

Homo sapiens.

Homo sapiens.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (Bases 1 to 1533)

Direct Submission

Strausberg, R.

Submitted (07-DEC-2001) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgapsb-f@mail.nih.gov

Tissue Procurement: DCTD/DP

cDNA Library Preparation: Rubin Laboratory

DNA Sequencing by: National Institutes of Health Intramural

Sequencing Center (NISC),

Gaithersburg, Maryland;

Web site: <http://www.nisc.nih.gov/>

Contact: nisc_mgc@nhgri.nih.gov

REMARK

COMMENT

SOURCE	Rattus norvegicus.	
ORGANISM	Rattus norvegicus	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;	
	Rattus.	
REFERENCE	1 (bases 1 to 1598)	
AUTHORS	Ciesla, J., Weiner, K.X., Weiner, R.S., Reston, J.T., Maley, G.F. and Maley, F.	
TITLE	Isolation and expression of rat thymidylate synthase cDNA:	
JOURNAL	phylogenetic comparison with human and mouse thymidylate synthases	
MEDLINE	Biochim. Biophys. Acta 1261 (2), 233-242 (1995)	
PUBMED	95226450	
FEATURES	7711067	
source	Location/Qualifiers	
	1..1598	
	/organism="Rattus norvegicus"	
	/strain="Reuber"	
	/db_xref="taxon:10116"	
	/cell_line="H35"	
	/cell_types="hepatoma"	
	1..924	
	/EC_number="2.1.1.45"	
	/function="conversion of dUMP to dTMP"	
	/note="dimer; putative"	
	/codon_start=1	
	/product="thymidylate synthase"	
	/protein_id="AAA92340.1"	
	/db_xref="GI:207328"	
	/translation="MLVEGSELQGAQOQRTAPQHGELVLRQVHIMRCGFKKEDR	
	TGTGILSVFGQARYSLRDEPPLITTKVFNKGLLELWFIKGSNNAKELSSKGVRI	
	WDANGSDFDLISGLFSARQGDGIVPGFQMRHFGADYKDMDSYQSGVDLQKVID	
	ITKNDPDRRLINCAWPKDLPLMPCHALCOFVNVNSELSCQIYKSGDMGLGVP	
	FNIAGYALLTVMIAHITGLQPGDFVHTLGDHAIYVLIHIEPIKIQLOREPRPFKRLIL	
	RKVTIDDFKVEDFQIRGYNPHPTIKEMAV"	
BASE COUNT	438 a 328 c 375 g 457 t	
ORIGIN		
	Query Match 45.4%; Score 697.8; DB 10; Length 1598;	
	Best Local Similarity 87.7%; Pred. No. 2.4e-134;	
	Matches 762; Conservative 0; Mismatches 107; Indels 0; Gaps 0;	
QY	181 CCGCCGACGCGGAGCTGCAGTACCTGGGCGAGATCCACACATCTCCGCTCGCGGCTC 240	
Db	58 CGCAGCATCGAGAACTCCAGTAACCTGAGGCAGGTGAGACATATTATGCGTTCGGCTTC 117	
QY	241 AGAAGAGACGCGACCGGACCGGACCTCTCGGTATTCGGCATGCGACGCGGCTAC 300	
Db	118 AAGAAGAGACCGGACCGGACCTCGGCTGCTGGCATGCGACGCGGTAC 177	
QY	301 AGCTGAGACATGAATTCCTCTGCTGACACCAACGCTGTCTTGGAGGCTGTTTG 360	
Db	178 AGCTGAGACATGAATTCCTCTGCTGACACCAACGAGTATTCGGAGGCTGTTTG 237	
QY	361 GAGGAGTGTGTTGTTTATCAAGGATCCAAATGCTAAAGAGCTGTCTTCCAAAGGA 420	
Db	238 GAGGAGTGTGTTGTTTATCAAGGCTCCAAATGCTAAAGAGTGTCTTCCAAAGGA 297	
QY	421 GTGAAATCTGGGATGCCAATGGATCCCGAGACTTTTGGACAGCCTGGGATTCCTACC 480	
Db	298 GTGAGATTTGGATGCCAATGGTCCCGAGACTTTTGGACAGCTGGGATTCCTGCC 357	
QY	481 AGAGAAGAAGGGGACTTGGGCCCGAGTTTATGGCTTCCAGTGGAGGCAATTTTGGGCGAGAA 540	
Db	358 CGACAGAGAGAGACTTGGGCCCGAGTTTATGGATTCAGTGGAGACATTTTGGAGCAGAC 417	
QY	541 TACAGAGATATGAATTCAGATATTTCAGAGACAGGAGTTGACCAATGCAAGAGTGTAT 600	
Db	418 TACAAAGATATGATTCAGATATTTCAGAGAGTGTAGACAGCTGCAAAAAGTGTAT 477	
QY	601 GACACCATCAAAACCAACCTCGACGACGAGAGATCATGTGCGCTGGGATCCAGA 660	
Db	478 GACACCATCAAAACCAACCTCGACGAGAGATCATGTGCTGCTGGAACCCAAA 537	

QY	661 GATCTTCCTCTGATGCGCGCTGCCTCCATGCATGCCTCTGCCAGTTCCTATGTGGTGAAC 720	
Db	538 GATCTTCCTCTGATGCGCGCTGCCTCCATGCATGCCTCTGTCATTTTATGTGGTGAAT 597	
QY	721 AGTCAGCTGTCTGCGCAGCTGTACAGAGATCGGGAGACATCGGCTCGGTGCTTC 780	
Db	598 GGGAGAGCTGTCTGCGCAGCTTTACAGGGTTCAGAGAGATATGGTCTGGGTGCGCCTTC 657	
QY	781 AACATCCGACGCTAGCGCCCTGCTCACTACATGATGATGGCAGATCAGCGGCTCGAAGCCA 840	
Db	658 AACATTGCCAGCTATGCTGCTGACCTACATGATGACATATCAGGGCTTCAGCGCG 717	
QY	841 GGTGACTTTATACACACTTTGGGAGATGCCACATATTACCTGAATCAGATCAGCCACTG 900	
Db	718 GGTGATTTTGTCCATACCTTTGGGAGATGCCACATTTATCTGAATCATATTGAGCCACTG 777	
QY	901 AAAATTCAGCTTCAGCGAGACCCAGACCTTTCCAAAGCTCAGGATTCCTTCGAAAAGTT 960	
Db	778 AAAATTCAGCTTCAGCGAGACCCAGACCTTTCCAAAGCTCAGGATTCCTTCGAAAAGTT 837	
QY	961 GAGAAATTCAGTACCTTCAAGCTCAGACCTTTTCAGATTGAAGGGTACAAATCCCATCCA 1020	
Db	838 GAGAAATTCAGTACCTTCAAGCTCAGACCTTTTCAGATTGAAGGGTACAAATCCCATCCA 897	
QY	1021 ACTATTAAATGGAATGCTGTTTGGG 1049	
Db	898 ACGATTAAATGGAATGCTGTTTGGG 926	

RESULT 15

MUSPSTSA

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

FEATURES

source

CDS

1349 bp

mRNA

linear

1349-1993

complete cds.

MUSPSTSA

Mouse thymidylate synthase pseudogene mRNA, complete cds.

M30774.1

GI:200561

pseudogene; thymidylate synthase.

Mouse J03-7 and 376 cell line DNA.

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 1349)

Li, D.W. and Johnson, L.F.

A mouse thymidylate synthase pseudogene derived from an aberrantly processed RNA molecule

Gene 82 (2), 363-370 (1989)

90060828

2573561

Location/Qualifiers

1..1349

/organism="Mus musculus"

/db_xref="taxon:10090"

177..1097

/note="pseudo-thymidylate synthase"

/pseudo

/codon_start=1

345 a 320 c 335 g 349 t

1 bp upstream of HindIII site.

Query Match 41.3%; Score 634.8; DB 10; Length 1349;

Best Local Similarity 84.6%; Pred. No. 2.8e-121;

Matches 767; Conservative 0; Mismatches 118; Indels 22; Gaps 4;

QY 181 CCGCCGACGCGGAGCTGCAGTACCTGGGCGAGATCCAAACATCTCCGCTCGCGGCTC 240 | || Db | 225 CCGCGGCATGGAGAACTCCAGTACCTGAGCGAGGTGAGCAGCATTTTGTGCTCGCGCTTC 284 | |
QY	241 AGAAGAGAGACCGGACCGGACCGGACCTCTGCTGATTCGGCATGCGAGCGGCTAC 300	
Db	285 AAGAAGAGAGACCGGACCGGACCTGTCACCTGTAGGTGTTCGGCATGCGAGCGGATCC 344	
QY	301 AGCTGAGAGATGAATTCCTCTGCTGACCAACGCTGTGTTCTGGAAGGCTGTTTG 360	
Db	345 AGCTGAGAGATGAATTCCTCTGCTGACCAACGCTGTGTTCTGGAAGGCTGTTTG 403	

```
QY 666 TCCTCTGATGGCGCTGCTCCATGCCATGCCCTCTGCCAGTCTTATGTGGTGAACAGTGA 725
Db 592 TCCCTGATGGCACTGCTCCCTTCCATGCCATGCCCTCTGCCAGTCTTATGTGGTGAACAGTGA 651
QY 726 GCTCTCTGCGCAGCTGTACACAGATCGGAGACATGGGCGCTGGTGGTCTTCAACAT 785
Db 652 ACTGCTTGGCAGCTTTACAGAGCTCAGAGATATGGGCTTGGGCGTGGCCTTCAACAT 711
QY 786 CGCAGCTACGCCCTCTCACTACATGATTCGACATCAATCGAGCTGAGCCAGTGA 845
Db 712 TGCCAGTATGCTCTGCTCACTACATGATTCGACATCAATCGAGCTGAGCCAGTGA 771
QY 846 CTTTATACACTTGGGAGATGACATATTTACTGATTCATGATTCAGCTGAGCCAGTGA 905
Db 772 TTTTGTCCACACTTGGGAGATGACATATTTACTGATTCATGATTCAGCTGAGCCAGTGA 831
QY 906 TCAGCTTCAGGAGACCCAGACCTTCCCAAGCTCAGATTTCTCGAAGAGTTGAGAA 965
Db 832 TCAGCTACAGGAGACCCAGACCTTCCCAAGCTCAGATTTCTCGAAGAGTTGAGAC 891
QY 966 AATTGATCACTTCAAGCTCAGAGCTTTTCAAGTGAAGGTTACAAATCCCAACTAT 1025
Db 892 AATCGATGATTTCAAGTTCAGAGCTTTTCAAGTGAAGGTTATATCCACATCCACAT 951
QY 1026 TAAATGCAATGGCTGCTTTA 1046
Db 952 TAAATGCAATGGCTGCTTTA 972

RESULT 13
MUSTHSM 961 bp mRNA linear ROD 27-APR-1993
LOCUS Mouse thymidylate synthase mRNA, complete cds.
DEFINITION
ACCESSION M13019
VERSION M13019.1 GI:202029
KEYWORDS thymidylate synthetase.
SOURCE Mouse, cDNA to mRNA, clone pMTS-3.
ORGANISM Mus musculus
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLES Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (Bases 1 to 961)
Perryman,S.M., Rossana,C., Deng,T.L., Vanin,E.F. and Johnson,L.F.
Sequence of a cDNA for mouse thymidylate synthase reveals striking
similarity with the prokaryotic enzyme
Mol. Biol. Evol. 3 (4), 313-321 (1986)
JOURNAL 88174353
MEDLINE 344407
PUBMED
COMMENT Draft entry and printed copy of the sequence [1] kindly provided by
L.F. Johnson, 29-AUG-1986.
FEATURES
source
1. .961
/organism="Mus musculus"
/db xref="taxon:10090"
<1. .961
/product="this mRNA"
38. .961
/notes="thymidylate synthase (EC 2.1.1.45)"
/codon start=1
/db xref="GI:202030"
/translation="MLVYSELSQDAQLSAPRHGELOYLRVHEILRCGPKEDR
TGUTLSVFMQARISLRDEPFLITTRVFKWGLVLELLFKGSTNAKELSKGRI
WDANGSEFDSIGSAPGEGLGVYQFWRHFGABYKMDSDYSGGVDDQLQVID
TITNPDRIITACWNPOLFMALPFCHALCOFYVNGELSCOLYQSGDMGLGV
FNIASVALLYMTAHTILOQDFVHTLGDHILYLNHIEPLIKILOREPRPPFKLIL
RKVETIDDFKVEDFQEGYNPHPTIKMEMAV"
BASE COUNT 253 a 224 c 251 g 233 t
ORIGIN 63 bp upstream of PstI site.
Query Match 46.0%; Score 706.6; DB 10; Length 961;
Best Local Similarity 85.5%; Pred. No. 3.4e-136;
Matches 787; Conservative 0; Mismatches 134; Indels 0; Gaps 0;
```

RESULT 14
RATTHYSYN
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS

RATTHYSYN 1598 bp mRNA linear ROD 30-MAY-2000
Rattus norvegicus thymidylate synthase mRNA, complete cds.
L12138
L12138.1 GI:207327
thymidylate synthase.

1026 TAAATGCAATGGCTGCTTTA 1046
940 TAAATGCAATGGCTGCTTTA 960

QY 126 GGTCCGCGCGCGCGCTTGTCCCGCCCGCCCGCCAGAGCGCGCGCGCTCGGCC 185
Db 40 GCTGTGTGTGGCTCCGAGCTGCACTCCGATGCTCAGCAGCTGAGCGCGGAAAGCCCGCAG 99
QY 186 GCACGGGAGCTGAGTACTTGGGGAGATCCACACATCTCTCGCTGGCGCTCAGGAA 245
Db 100 GCATGGAGAACTCCAGTACCTGAGGAGGCTGGAGCACTTTTGGCTGGCGCTTCAAGAA 159
QY 246 GCACGACCGCAGCGGACCGGACCGCTGTTCGGTATTCGGATCGAGCGCGCTACAGCT 305
Db 160 GGAGGACCGCAGCGGACCGGACCGCTGTTCGGTATTCGGATCGAGCGCGCTACAGCT 219
QY 306 GAGAGATCAATTCCTCTGCTGACAAACCAAGCTGTCTTGGAGAGGCTTTTGGAGGA 365
Db 220 GAGAGATCAATTCCTCTGCTGACAAACCAAGCTGTCTTGGAGAGGCTTTTGGAGGA 279
QY 366 GTTGTGTGTGTATCAAGGATCCCAAAATCTTAAAGAGCTGTCTTCCAAGGAGTGA 425
Db 280 GTTGTGTGTGTATCAAGGATCCCAAAATCTTAAAGAGCTGTCTTCCAAGGAGTGA 339
QY 426 AATCTGGGATGCCAATGATTCGAGACTTTTGGACAGCTGGGATTTTCCACAGAGA 485
Db 340 AATCTGGGATGCCAATGATTCGAGACTTTTGGACAGCTGGGATTTTCCACAGAGA 399
QY 486 AGAGGGGACTTGGGCGCCAGTTTATGCTTCCAGTGGAGGATTTTGGGCGAGATACAG 545
Db 400 GGAAGGGGACCTGGGCGCCAGTTTATGCTTCCAAATGGAGGCAATTTTGGAGCAGATACAA 459
QY 546 AGATATGGAATCAGATTTTACGACAGGAGGCTTCAACCACTGCAAGAGATGATTCACAC 605
Db 460 AGATATGGAATCAGATTTTACGACAGGAGGCTTCAACCACTGCAAGAGATGATTCACAC 519
QY 606 CATCAAAACCAACCTCAACGACAGAGGATTCATCATGTGCGCTTGGAAATCCAGAGATCT 665
Db 520 CATCAAAACCAACCTCAACGACAGAGGATTCATCATGTGCGCTTGGAAATCCAGAGATCT 579
QY 666 TCTCTGTATGCGCTGCTCCATGCCATGCCCTCTGCCAGATTTCTATGTGTGAACAGTGA 725
Db 580 TCCCTGTATGCGCTGCTCCATGCCATGCCCTCTGTCTGTCTATGTGTGAATGGGA 639
QY 726 GCTGTCTGCGCAGCTGTACACAGATCGGAGACATGGGCGCTGGTGTCTTCAACAT 785
Db 640 ACTGTCTTGGCAGCTTTTACAGAGGCTCAGAGATATGGTCTTGGGCGTGGCTTCAACAT 699
QY 786 CGCAGCTACGCGCTGCTCAGTATGATTTGCGACATCAGCGGCTGAGCGCAGTGA 845
Db 700 TGCCAGCTATGCTCTGCTCACCTACATGATTCACATATCAACAGGCTTGCAGCCAGTGA 759
QY 846 CTTTATACACATTTGGGAGATCAGATTTTACCTGATTCACATCGAGCGCTGAAAT 905
Db 760 TTTTGTCCACACTTTGGGAGATCAGATTTTACCTGATTCACATCGAGCGCTGAAAT 819
QY 906 TCAGCTTCAGCGAGAACCCAGACCTTTCCCAAGCTCAGGATTTCTCGAAAAGTTGAGAA 965
Db 820 TCAGCTTCAGCGAGAACCCAGACCTTTCCCAAGCTCAGGATTTCTCGAAAAGTTGAGAC 879
QY 966 AATTGATCACTTCAAGCTCAGAGCTTTTCAAGTGAAGGTTACAAATCCGATCCAACTAT 1025
Db 880 AATCGATGATTTCAAGCTTGAAGCTTTTCAAGTGAAGGTTATATCCATCCACATCCAGT 939
QY 1026 TAAATGCAATGGCTGCTTTA 1046
Db 940 TAAATGCAATGGCTGCTTTA 960

Richard S. Gibbs, R.A.

Richards, S., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/ILLUMI at: <http://image.llnl.gov>
Series: IRAP Plate: 36 Row: J Column: 13
This clone was selected for full length sequencing because it
passed the following selection criteria: Similarity but not
passivity to protein.

FEATURES source

```

1. 986
/organism="Mus musculus"
/db_xref="taxon:10090"
/map="CZECH 11"
/clone="MGC:28246 IMAGE:3994204"
/tissue type="mammary tumor metastasized to lung. Tumor
arose spontaneously from a senescent normal mammary
(cloneal) outgrowth infected with the virus MMTV."
/clone_lib="NCI CGAP_Lu29"
/lab host="DH10B"
/notes="Vector: pCWB-SPORT6"
50. .973
/codon start=1
/product="Unknown (protein for MGC:28246)"
/protein_id="AAH20139.1"
/db_xref="GI:18044897"
/translation="MLVGSFELSDAQQLSAEPQHGELQYLRVEHLISGCKFKEDR
TGCTLTSLFQNGARYSARDEDEPPLTTKRPYKGVLEELLTFKGSNAGELSSKGWRI
WDANGSRDLYSQAQDEGDIYPGPQWPHFGAYYKMGDSYSGGVQDLQKVID
TKINPDRELRICMAWNPDLPLMAPLCPFYVNGMSLCQLYRSGDMGLGVP
FNIA5YALLFMYHTALTGLQSGDFVTLGDGAILYLNHLEPLQLQRPFRPFPLKIL
RKVETLIDDFKVEQDFQEGNVPHTIKEMAW"
237 t

```

505

[illegible]

RESULT 12	986 bp	linear	complete cds.	ROD 07-AUG-2002
EC020139	mRNA			

BC020139	986 bp	10000	complete cds.
LOCUS			
DEFINITION	Mus musculus, clone MGC:28246 IMAGE:3994204, mRNA, complete cds.		
ACCESSION	BC020139		
VERSION	BC020139.1	GI:18044896	
KEYWORDS	MGC.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus		
	1 (bases 1 to 986)		
REFERENCE	Strausberg, R.		
AUTHORS	Direct Submission		
TITLE	Submitted (19-DEC-2001) National Institutes of Health, Mammalian		
JOURNAL	Gene Collection (MGC), Cancer Genomics Office, National Cancer		
	Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,		
	USA		
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov		
COMMENT	Contact: MGC help desk		
	Email: cgapbs-remail.nih.gov		
	Tissue Procurement: Gilbert Smith, Ph.D.		
	CNA Library Preparation: Life Technologies, Inc.		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)		
	DNA Sequencing by: Baylor College of Medicine Human Genome		
	Sequencing Center		
	Center code: BCM-HGSC		
	Web site: http://www.hgsc.bcm.tmc.edu/cdna/		
	Contact: amg@bcm.tmc.edu		
	Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,		
	Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,		

KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
gene
CDS
BASE COUNT
ORIGIN
Query Match
Best Local Similarity
Matches 840; Conservative 0; Mismatches 0; Indels 102; Gaps 1;
QY 106 ATGCCTGTGCCGGTCGAGAGTCGCGCGCGCGCCCTTGTCCCGCGCGCACAGGAGCGG 165
Db 1 ATGCCTGTGCCGGTCGAGAGTCGCGCGCGCGCCCTTGTCCCGCGCGCACAGGAGCGG 60
QY 166 GAGCCGAGCGCGTCGCCCGCACGGGAGCTGCGAGTACTTGGGGCAGATCCAAACATC 225
Db 61 GAGCGCGAGCGCGTCGCCCGCACGGGAGCTGCGAGTACTTGGGGCAGATCCAAACATC 120
QY 226 CTCGCTGCGCGCTCAGAGAGGAGCCGACGGGACCCGACCCCTGTGGGTATCGGC 285
Db 121 CTCGCTGCGCGCTCAGAGAGGAGCCGACGGGACCCGACCCCTGTGGGTATCGGC 180
QY 286 ATGACGGCGCGCTACAGCCTGAGAGATGAATTCCTCTGTGTGCAACCAAGAGCTGTGTTC 345
Db 181 ATGACGGCGCTACAGCCTGAGAGATGAATTCCTCTGTGTGCAACCAAGAGCTGTGTTC 240
QY 346 TGAAGCGGTGTTTGGAGGAGTTCGTGTGTTTATCAAGGGATCCCAATGCTAAAGAG 405
Db 241 TGAAGCGGTGTTTGGAGGAGTTCGTGTGTTTATCAAGGGATCCCAATGCTAAAGAG 300
QY 406 CTGCTTCCAGGAGTGAAAATCTTGGGATGCCAATGGATCCGAGACTTTTGGACAGC 465
Db 301 CTGCTTCCAGGAGTGAAAATCTTGGGATGCCAATGGATCCGAGACTTTTGGACAGC 360

QY	1230	TAATAAAGGCTTTGAGTTAACTCACTGAGGGTATCTGACAAATGCTGAGGTATGAACAA	1289
Db	1206	TAAATAAAGGCTTTGAGTTAACTCCCTGAGGGTATCTGACAAATGCTGAGGTATGAACAA	1265
QY	1290	AGTCGAGGAGTAAGTAATGATGCTCTTTAGCAAAAACATGATGATGTCATTTCAATCCCA	1349
Db	1266	AGTCGAGGAGTAAGTAATGATGCTCTTTAGCAAAAACATGATGATGTCATTTCAATCCCA	1325
QY	1350	CGTACTTATAAAGAGGTTGGTGAATTTCAAGAGCTATTTTGGAAATATTTTGAATAT	1409
Db	1326	CGTACTTATAAAGAGGTTGGTGAATTTCCCAAGCTATTTTGGAAATATTTTGAATAT	1385
QY	1410	TTTAAAGAAATTTCAAGCTATTTCCCTCAATCTGAGGAGCTGAGTAACACCATCGATCA	1469
Db	1386	TTTAAAGAAATTTCCCAAGCTATTTCCCTCAATTTGAGGAGCTGAGTAACCCATCGATCA	1445
QY	1470	TCATGTAGAGTGTGGTTAAGAAGCT-----TTAAGTTGTTTTATATGTTGCTATAATAA	1523
Db	1446	TCATGTAGAGTGTGGTTAAGAAGCTTTAAAGTTATAGTTGTTTTATATGTTGCTATAATAA	1505
QY	1524	AGAAGTTTCTGC	1536
Db	1506	AGAAGTGTTTTGC	1518
RESULT 9			
LOCUS	AX069365	942 bp	DNA linear PAT 25-JAN-2001
DEFINITION	Sequence 29 from Patent WO0102600.		
ACCESSION	AX069365		
VERSION	AX069365.1	GI:12579227	
KEYWORDS	human.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	Yuan,C.S.		
JOURNAL	Detection of analytes using attenuated enzymes		
	Patent: WO 0102600-A 29 11-JAN-2001;		
FEATURES	GENERAL ATOMICS (US)		
source	Location/Qualifiers		
	1..942		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/note="Human thymidylate synthase gene: exons 1-8"		
BASE COUNT	236 a 247 c 257 g 202 t		
ORIGIN			
Query Match	61.3%; Score 942; DB 6; Length 942;		
Best Local Similarity	100.0%; Pred. No. 5e-185;		
Matches 942; Conservative	0; Mismatches 0; Indels 0; Gaps 0;		
QY	106	ATGCTCTGTGGCGGGCTCGGAGCTGCCGCGCGGCTTCGCCCGCCCGCCGACAGGAGCGG	165
Db	1	ATGCTCTGTGGCGGGCTCGGAGCTGCCGCGCGGCTTCGCCCGCCCGCCGACAGGAGCGG	60
QY	166	GACGCGGAGCGCGGTCCGCGCAGCGGGAGCTGAGTAAGTGGGGCAGATCCACACATC	225
Db	61	GACGCGGAGCGCGGTCCGCGCAGCGGGAGCTGAGTAAGTGGGGCAGATCCACACATC	120
QY	226	CTCCGCTCGGCGTTCAGGAAGACGACCGGACCGGACCGGACCCCTTCGGTATTCGGC	285
Db	121	CTCCGCTCGGCGTTCAGGAAGACGACCGGACCGGACCGGACCCCTTCGGTATTCGGC	180
QY	286	ATGCGAGGCGCTACAGCCTTGAGATGAATTCCTCTGCTGACCAACCAACGCTGCTTC	345
Db	181	ATGCGAGGCGCTACAGCCTTGAGATGAATTCCTCTGCTGACCAACCAACGCTGCTTC	240
QY	346	TGGAAGGGTGTTTTGGAGGATGTCGTGGTTTATCAAGGATTCACAAATGCTAAAGAG	405
Db	241	TGGAAGGGTGTTTTGGAGGATGTCGTGGTTTATCAAGGATTCACAAATGCTAAAGAG	300

```

Db 1340 CGTACTTATTAAGAGGTTGGTGAATTTTCAACAGCTATTTTGGAAATATTTTGAATAT 1399
QY 1410 TTTAAGAAATTCACAGCTATTTCCCTCAAATCTGAGGGAGCTGAGTAACACATCGATCA 1469
Db 1400 TTTAAGAAATTCACAGCTATTTCCCTCAAATCTGAGGGAGCTGAGTAACACATCGATCA 1459
QY 1470 TGATGTAGAGTGTGGTATGAACCT-----TTANAGTTGTTTATATATGTGCTATAATAA 1523
Db 1460 TGATGTAGAGTGTGGTATGAACCTTTAAAGTTATAGTTGTTTATATATGTGCTATAATAA 1519
QY 1524 AGAAGTGTCTGC 1536
Db 1520 AGAAGTGTCTGC 1532

RESULT 8
AB062290 1567 bp mRNA linear PRI 23-MAY-2002
LOCUS Homo sapiens. OK/SW-cl.29 mRNA for thymidylate synthetase, complete
DEFINITION Homo sapiens. OK/SW-cl.29 mRNA for thymidylate synthetase, complete
cds.
ACCESSION AB062290
VERSION AB062290.1 GI:21104405
KEYWORDS
SOURCE
ORGANISM Homo sapiens
            Homo sapiens colon adenocarcinoma cell_line:SW620 cDNA to mRNA.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Shichijo,S. and Itoh,K.
TITLE Identification of immuno-peptidmics that recognized by
        tumor-reactive CTL generated from TIL of colon cancer patients
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1567)
AUTHORS Itoh,K. and Shichijo,S.
TITLE Direct Submission
JOURNAL Submitted (24-MAY-2001) Shigeki Shichijo, Kurume University School
        of Medicine, Department of Immunology; 67 Asahi-machi, Kurume,
        Fukuoka 830-0011, Japan (E-mail:shichijo@med.kurume-u.ac.jp,
        Tel:81-942-31-7551, Fax:81-942-31-7699)
FEATURES
            Location/Qualifiers
            1..1567
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /cell_line="SW620"
                /cell_type="colon adenocarcinoma"
            1..1567
                /gene="OK/SW-cl.29"
            82..1023
                /gene="OK/SW-cl.29"
                /codon_start=1
                /product="thymidylate synthetase"
                /protein_id="BAB93473.1"
                /db_xref="GI:21104406"
                /translation="MPVAGSELPRRLPFPAAQERDAEPHPGELQYLQIQHILRGG
                VRKDRGTGSLVFCMOARYSLRDEPLLTTRVFWKGLVLELWFLKGSTNAKELS
                SKGVKWDANGSDFFDLSLGSREBDLGVFQWRFHGAERYDMSDYSGGVDO
                LQRVDTIKNPDRRLINMCANPRDLPALPPCHALQCFYVNSELSCOLYORSQ
                MGLGVFNPIASVALLYTAHTITGLKPGDFITLGDHLYLHNLHPKIQIQLQREPRFF
                PKLRILKVEKIDDFKAEQDFQEGYNPHPTIKEMAV"
BASE COUNT 431 a 366 c 389 g 381 t
ORIGIN
Query Match 95.9%; Score 1473.6; DB 9; Length 1567;
Best Local Similarity 98.8%; Pred. No. 2.7e-295;
Matches 1495; Conservative 0; Mismatches 12; Indels 6; Gaps 1;

QY 30 CGTCCGCGCGCCACTTGGCGCTGCCTCCGTCGCGCGGCACATTGGCTGCCTCGCTC 89
Db 6 CCTGCCACCGCCACTTGGCGCTGCCTCCGTCGCGCGGCACATTGGCTGCCTCGCTC 65
QY 90 CCGCGCGCGCGCCACTTGGCGCTGCCTCCGTCGCGCGGCACATTGGCTGCCTCGCTC 149
Db 66 CCGCGCGCGCGCCACTTGGCGCTGCCTCCGTCGCGCGGCACATTGGCTGCCTCGCTC 125
```

```

150 CCGCGCACAGGAGCGGACCGCCGAGCGCGTCCGCCGACGCGGAGCTGCAGTACCTGGG 209
126 CCGCGCACAGGAGCGGACCGCCGAGCGCGTCCGCCGACGCGGAGCTGCAGTACCTGGG 185
210 GCAGATCCAAACAATCTCTCCGCTGCGCGCTGAGGAAGACGACCGGACCGGAC 269
186 GCAGATCCAAACAATCTCTCCGCTGCGCGCTGAGGAAGACGACCGGACCGGAC 245
270 CTTGTCGGTATTCGGCATTCGAGCGCGCTACAGCTGAGAGATGAATTCCTCTGCTGAC 329
246 CTTGTCGGTATTCGGCATTCGAGCGCGCTACAGCTGAGAGATGAATTCCTCTGCTGAC 305
330 AACCAAAAGTGTCTGGAAGGCTGTTTTCGAGGAGTGTCTGTGTTTATCAAGGATC 389
306 AACCAAAAGTGTCTGGAAGGCTGTTTTCGAGGAGTGTCTGTGTTTATCAAGGATC 365
390 CACAAATGCTAAAGAGCTGCTTCCAAAGGAGTGAAATCTGGGATGCCAAATGATCCCG 449
366 CACAAATGCTAAAGAGCTGCTTCCAAAGGAGTGAAATCTGGGATGCCAAATGATCCCG 425
450 AGACTTTTTCGACAGCCTTGGGATTTCTCCACGAGAGAAGAGGGGACTTTGGGCCCACTTTA 509
426 AGACTTTTTCGACAGCCTTGGGATTTCTCCACGAGAGAAGAGGGGACTTTGGGCCCACTTTA 485
510 TGGCTTTCAGTGGAGGCAATTTTGGGCGACAATACAGAGATATGGAATCAGATATTTCAGG 569
486 TGGCTTTCAGTGGAGGCAATTTTGGGCGACAATACAGAGATATGGAATCAGATATTTCAGG 545
570 ACAGGAGTTCACAACTGCAAAAGAGTGATTCGACACCATCAAAAACCAACCTGACACAG 629
546 ACAGGAGTTCACAACTGCAAAAGAGTGATTCGACACCATCAAAAACCAACCTGACACAG 605
630 AAGATATCATATGTCGCTTGGATTCGAATCCAGAGATCTTCTCTGATGGCGCTGCTCCATG 689
606 AAGATATCATATGTCGCTTGGATTCGAATCCAGAGATCTTCTCTGATGGCGCTGCTCCATG 665
690 CCATGCGCTCTGCCAGTTCTATGTTGGTGAACAGTGAAGTGTCTGCGCAGCTGTACACAGAG 749
666 CCATGCGCTCTGCCAGTTCTATGTTGGTGAACAGTGAAGTGTCTGCGCAGCTGTACACAGAG 725
750 ATCGGAGACATGCGGCTCGGCTGTCCTTTCACATTCGCGAGCTACGCGCTGCTCAGTA 809
726 ATCGGAGACATGCGGCTCGGCTGTCCTTTCACATTCGCGAGCTACGCGCTGCTCAGTA 785
810 CATGATTCGCGACATTCACGCGCGCTGGAAGCAGGTGACTTTATACACTTTTGGGAGATGC 869
786 CATGATTCGCGACATTCACGCGCGCTGGAAGCAGGTGACTTTATACACTTTTGGGAGATGC 845
870 ACATATTTACCTGGAATTCACATTCGAGCCACTGAAATTCAGCTTCAGCGAGAACCCAGACC 929
846 ACATATTTACCTGGAATTCACATTCGAGCCACTGAAATTCAGCTTCAGCGAGAACCCAGACC 905
930 TTTCCCAAGAGCTCAGGATCTTTCGAAAGTTCGAAATTCAGTACCTTCAAGCTCAAGA 989
906 TTTCCCAAGAGCTCAGGATCTTTCGAAAGTTCGAAATTCAGTACCTTCAAGCTCAAGA 965
990 CTTTTCAGATTGAAGGTACAACTCCGATCAACTTAAATTTAAATTTAAATTTAAATTTAAATTT 1049
966 CTTTTCAGATTGAAGGTACAACTCCGATCAACTTAAATTTAAATTTAAATTTAAATTTAAATTT 1025
1050 TGCCTTTCAAGAGAGCTGCAAGGATATTTGTCAGTCTTTAGGGGTTTGGGCTCGATCCCGAGG 1109
1026 TGCCTTTCAAGAGAGCTGCAAGGATATTTGTCAGTCTTTAGGGGTTTGGGCTCGATCCCGAGG 1085
1110 TAAAGTTCTTTTTCCTCTAAAGANAAGGAACTAGTGTCAAAAATCTGTCCGTGACCTTA 1169
1086 TAAAGTTCTTTTTCCTCTAAAGANAAGGAACTAGTGTCAAAAATCTGTCCGTGACCTTA 1145
1170 TCAGTTATTAATTTTAAAGGATGTTGCCACTGGCAAAATGTAACCTGTCCAGTCTTTTCCA 1229
1146 TCAGTTATTAATTTTAAAGGATGTTGCCACTGGCAAAATGTAACCTGTCCAGTCTTTTCCA 1205
```


SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1569)
Direct Submission
Strausberg, R.
Submitted (07-SEP-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapsb-re@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
contact: amandansystemsbiology.org
Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia
Greene, Mark Kettman and Anuradha Madan

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate; 31 Row; 8 Column; 4
This clone was selected for full length sequencing because it
passed the following selection criteria: Similarity but not
identity to protein.
Location/Qualifiers
1..1569
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:22884 IMAGE:4048625"
/tissue type="Bone marrow, chronic myelogenous leukemia"
/clone_lib="NIH_MGC_54"
/lab host="DH10B"
/note="Vector: pDNR-LIB"
96..1037
/codon_start=1
/product="Similar to thymidylate synthetase"
/protein_id="AAH1391.1"
/db_xref="GI:15530267"
/translation="MPVASELP...RRPLPAAQERDARPPHGELOYLGOIHILRCG
VRKDRGTGTLVFGNQARYSLRDEFFLLTKVFWKVLLELLWFKIGSTNAKELS
SKGVKWDANGSRDLDSFGSTREGBDLFPVQWRHFGARYMESDVGQGVDO
LQRVIDTKNPDRRIINCAWNPRLDLPALCPYVNVSELSCQYQKSGD
MGLVPPNFIASVALLTHIAHTGLKPGDFHTLGLDAHYLNHIEPLKIQLOREPRP
PKLRLLRKEKIDDKAEDFQIEGVNPHPTIKMEMAV"
BASE COUNT 425 a 368 c 398 g 378 t
ORIGIN

Query Match 96.6%; Score 1483.2; DB 9; Length 1569;
Best Local Similarity 99.2%; Pred. No. 2.8e-297;
Matches 1501; Conservative 0; Mismatches 6; Indels 6; Gaps 1;
QY 30 CGTCCGCGCGCCACTTGGCTGCTCGCTCCGCGCGCCACTTGGCTCCGCTCCGTC 89
Db 20 CCGCCACCGCGCCACTTGGCTGCTCGCTCCGCGCGCCACTTGGCTCCGCTCCGTC 79
QY 90 CCGCGCGCGCGCCACTTGGCTGCTCGCTCCGCGCGCCACTTGGCTCCGCTCCGTC 149
Db 80 CCGCGCGCGCGCCACTTGGCTGCTCGCTCCGCGCGCCACTTGGCTCCGCTCCGTC 139
QY 150 CGCGGACAGGAGCGGAGCGCGCGCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCG 209
Db 140 CGCGGACAGGAGCGGAGCGCGCGCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCG 199
QY 210 GCAGATCCACACATCTCCGCTGGCGGCTCAGGAGGCGCGCGCGCGCGCGCGCGCGCG 269
Db 200 GCAGATCCACACATCTCCGCTGGCGGCTCAGGAGGCGCGCGCGCGCGCGCGCGCGCG 259
QY 270 CCTGTGGTATTCGGCATCGAGCGCGCGCTACAGCTCAGAGATGAATTCCTCTGCTGAC 329

Db 260 CCTGTGGTATTCGGCATCGAGCGCGCTACAGCTCAGAGATGAATTCCTCTGCTGAC 319
QY 330 AACCAACCGTGTGTTCTGGAAGGGTGTGTTGGAGGAGTGTGTTGTTATCAAGGATC 389
Db 320 AACCAACCGTGTGTTCTGGAAGGGTGTGTTGGAGGAGTGTGTTGTTATCAAGGATC 379
QY 390 CACAAATGCTAAAGAGCTGTCTTCCAAAGGAGTGAATACTGGGATGCCAATGCATCCG 449
Db 380 CACAAATGCTAAAGAGCTGTCTTCCAAAGGAGTGAATACTGGGATGCCAATGCATCCG 439
QY 450 AGACTTTTGGACAGCGCTGGATCTCCACAGAGAGAGAGGAGCTTGGGCCAGTTTA 509
Db 440 AGACTTTTGGACAGCGCTGGATCTCCACAGAGAGAGAGGAGCTTGGGCCAGTTTA 499
QY 510 TGGCTTCCAGTGGAGCAATTTGGGGCAGAAATACAGAGATATGGAATCAGATTATTCAG 569
Db 500 TGGCTTCCAGTGGAGCAATTTGGGGCAGAAATACAGAGATATGGAATCAGATTATTCAG 559
QY 570 ACAGGAGTTCACCACTGCAAGAGTGAATGACCAATCAAAACCAACCTGAGCAG 629
Db 560 ACAGGAGTTCACCACTGCAAGAGTGAATGACCAATCAAAACCAACCTGAGCAG 619
QY 630 AAGATCATCATGTGGCTTGAATCCAAAGAGATCTTCTGATGGCGCTGCTCCATG 589
Db 620 AAGATCATCATGTGGCTTGAATCCAAAGAGATCTTCTGATGGCGCTGCTCCATG 679
QY 690 CCATGCCCTCTGCCAGTCTATGTGTTGAACAGTGTGCTCCAGAGCTGTACAGAG 749
Db 680 CCATGCCCTCTGCCAGTCTATGTGTTGAACAGTGTGCTCCAGAGCTGTACAGAG 739
QY 750 ATCGGGAGCATGGCGCTGGGTGCTTTCAACATCGCAGCTACGCCCTGCTCAGTA 809
Db 740 ATCGGGAGCATGGCGCTGGGTGCTTTTAAACATCGCAGCTACGCCCTGCTCAGTA 799
QY 810 CATGATTTGGCACAATCAAGGAGTGAATGACCAATCAAAACCAACCTGAGCAG 869
Db 800 CATGATTTGGCACAATCAAGGAGTGAATGACCAATCAAAACCAACCTGAGCAG 859
QY 870 ACATATTTACCTGAATCAGATCGACCACTGAAATTCAGCTTCAGCGAGAACCCAG 929
Db 860 ACATATTTACCTGAATCAGATCGACCACTGAAATTCAGCTTCAGCGAGAACCCAG 919
QY 930 TTTCCCAAGCTCAGGATCTTCGAAAGTGAATAATGATGATCTCAAGCTGAAGA 989
Db 920 TTTCCCAAGCTCAGGATCTTCGAAAGTGAATAATGATGATCTCAAGCTGAAGA 979
QY 990 CTTTCAGATTGAAGGTACATCCCGATCCCACTTAAATGAAATGGAATGGCTGTAGG 1049
Db 980 CTTTCAGATTGAAGGTACATCCCGATCCCACTTAAATGAAATGGAATGGCTGTAGG 1039
QY 1050 TGCCTTCAAGAGGCTGAGGATATGTCAGTCTTTAGGGTTGGGCTGGATGCCAGG 1109
Db 1040 TGCCTTCAAGAGGCTGAGGATATGTCAGTCTTTAGGGTTGGGCTGGATGCCAGG 1099
QY 1110 TAAAGATTTCTTTTGTCTTAAAGAAAGAACTAGGTCAAAATCTGTCGGTGACTA 1169
Db 1100 TAAAGATTTCTTTTGTCTTAAAGAAAGAAAGAACTAGGTCAAAATCTGTCGGTGACTA 1159
QY 1170 TCAGTTATTAATTTTAAAGATGTTGCCATCGGCAATGTAATGTCGCGAGTTCTTCCA 1229
Db 1160 TCAGTTATTAATTTTAAAGATGTTGCCATCGGCAATGTAATGTCGCGAGTTCTTCCA 1219
QY 1230 TAAATAAGGCTTTGATGTTAACTCACTGAGGTTATCTGCAATGCTGAGTTATGAACA 1289
Db 1220 TAAATAAGGCTTTGATGTTAACTCACTGAGGTTATCTGCAATGCTGAGTTATGAACA 1279
QY 1290 AGTGAAGAGAAATCAAAATGATGCTCTTACCAAAACATGATGATGCAATTTCAATCCA 1349
Db 1280 AGTGAAGAGAAATCAAAATGATGCTCTTACCAAAACATGATGATGCAATTTCAATCCA 1339
QY 1350 CCGTATTAAGAGGTTGGTGAATTTCAACAGCTATTTTTCGAAATATTTTAGAATAT 1409

```

Db 1508 c 1508

RESULT 6
AX281791
LOCUS
DEFINITION Sequence 200 from Patent WO0177389.
ACCESSION AX281791
VERSION AX281791.1 GI:16609042
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Shifman,D., Somogyi,R., Lawn,R., Seilhamer,J.J., Porter,G.J.,
Mikita,T. and Tai,J.
TITLE Genes expressed in foam cell differentiation
JOURNAL Patent: WO 0177389-A 200 18-OCT-2001;
Incyte Genomics, Inc. (US)
FEATURES
source
1..3298
/organism="Homo sapiens"
/db xref="taxon:9606"
/notes="Incyte ID No: 247220.15"
BASE COUNT 898 a 709 c 744 g 919 t 28 others
ORIGIN
1;
Query Match 97.2%; Score 1493.4; DB 6; Length 3298;
Best Local Similarity 99.3%; Pred. No. 2.3e-299;
Matches 1510; Conservative 0; Mismatches 5; Indels 6; Gaps 1;

QY 22 CTGCTCCGTCGCGCGCGCATTGGCTCCCTCCGTCGCGCGCGCATTGGCTG 81
Db 1 CTTGCTCCGTCGCGCGCGCATTGGCTCCCTCCGTCGCGCGCGCATTGGCTG 60
QY 82 CTTGCTCCGTCGCGCGCGCATTGGCTCCCTCCGTCGCGCGCGCATTGGCTG 141
Db 61 CTTGCTCCGTCGCGCGCGCATTGGCTCCCTCCGTCGCGCGCGCATTGGCTG 120
QY 142 TTGCTCCGTCGCGCGCGCATTGGCTCCCTCCGTCGCGCGCGCATTGGCTG 201
Db 121 TTGCTCCGTCGCGCGCGCATTGGCTCCCTCCGTCGCGCGCGCATTGGCTG 180
QY 202 TACCTGGGCGAGATCAACATCTCTCCGTCGCGCGCGCATTGGCTCCCTCCG 261
Db 181 TACCTGGGCGAGATCCAAACATCTCTCCGTCGCGCGCGCATTGGCTCCCTCCG 240
QY 262 ACCGCGACCTGTGCGTATTGGGATGCGGCGCGCATTGGCTCCCTCCGTCGCG 321
Db 241 ACCGCGACCTGTGCGTATTGGGATGCGGCGCGCATTGGCTCCCTCCGTCGCG 300
QY 322 CTGCTGACAAACAAACGTTGTTCTGGAAGGTTGTTGGAGGATGTTGGTTATC 381
Db 301 CTGCTGACAAACAAACGTTGTTCTGGAAGGTTGTTGGAGGATGTTGGTTATC 360
QY 382 AAGGATCCAAATGCTAAAGAGCTGTTCTCAAGGAGTCAAAATCTGGATGCCAAT 441
Db 361 AAGGATCCAAATGCTAAAGAGCTGTTCTCAAGGAGTCAAAATCTGGATGCCAAT 420
QY 442 GGATCCGAGATTTTGGACGCTGGGATCTCCACAGAGAGAGAGGAGCTTGGGC 501
Db 421 GGATCCGAGATTTTGGACGCTGGGATCTCCACAGAGAGAGAGGAGCTTGGGC 480
QY 502 CAGTTTATGCTTCCAGTGGAGCATTGTTGGGCGAGATATCAGATATGGAATCAGAT 561
Db 481 CAGTTTATGCTTCCAGTGGAGCATTGTTGGGCGAGATATCAGATATGGAATCAGAT 540
QY 562 TATTGAGGACGAGGAGTTGACCACTGCAAGAGTGAATGACCATCAAAACCAACCT 621
Db 541 TATTGAGGACGAGGAGTTGACCACTGCAAGAGTGAATGACCATCAAAACCAACCT 600

```

```

QY 622 GACGACAGAGATCATCATGTGGCTTGGAAATCCAGAGATCTTCTCTGATGGCGTG 681
Db 601 GACGACAGAGATCATCATGTGGCTTGGAAATCCAGAGATCTTCTCTGATGGCGTG 660
QY 682 CTTCCATGCGATGCGCTCTGCGAGTTCTATGTGTGTGAACAGTGAAGTGTCTCTG 741
Db 661 CTTCCATGCGATGCGCTCTGCGAGTTCTATGTGTGTGAACAGTGAAGTGTCTCTG 720
QY 742 TACCAGAGATCGGAGACATGGCGCTCGGTGTGCTTCAACATCGCCAGCTACGCCCTG 801
Db 721 TACCAGAGATCGGAGACATGGCGCTCGGTGTGCTTCAACATCGCCAGCTACGCCCTG 780
QY 802 CTCACGTACATGATTGGCGCATCATCGCGCTCAAGCCAGGTGACTTTTATACACATTG 861
Db 781 CTCACGTACATGATTGGCGCATCATCGCGCTCAAGCCAGGTGACTTTTATACACATTG 840
QY 862 GGAGATGACATATTTACCTGAATCATCATCGAGCCACTGAAAAATTCAGCTTCAGCGAGAA 921
Db 841 GGAGATGACATATTTACCTGAATCATCATCGAGCCACTGAAAAATTCAGCTTCAGCGAGAA 900
QY 922 CCAGACCTTTCCAAAGCTCAGATTTCTTCAAAAAGTTGAGAAAAATGATGACTTCAA 981
Db 901 CCAGACCTTTCCAAAGCTCAGATTTCTTCAAAAAGTTGAGAAAAATGATGACTTCAA 960
QY 982 GCTGAAGACTTTCAGATTGAAGGTGACATCCGATCCAACTATTTAAATGGAATGGCT 1041
Db 961 GCTGAAGACTTTCAGATTGAAGGTGACATCCGATCCAACTATTTAAATGGAATGGCT 1020
QY 1042 GTTTAGGTGCTTTCAAGGAGCTNGAAGATATTGTCAAGTCTTTAGGGGTGGCGTGA 1101
Db 1021 GTTTAGGTGCTTTCAAGGAGCTCGAAGATATTGTCAAGTCTTTAGGGGTGGCGTGA 1080
QY 1102 TGCCGAGGTAAAGTCTTTTCTCTTAAAGAAAGAACTAGGTCAAAAATCTGTCC 1161
Db 1081 TGCCGAGGTAAAGTCTTTTCTCTTAAAGAAAGAACTAGGTCAAAAATCTGTCC 1140
QY 1162 GTGACCTATCAGTTATTAAATTTTAAAGATGTTGCCACTGGCAATGTAATGTGCGAGT 1221
Db 1141 GTGACCTATCAGTTATTAAATTTTAAAGATGTTGCCACTGGCAATGTAATGTGCGAGT 1200
QY 1222 TCTTTCATATAAAGCTTTGAGTTAACTCACTGAGGATATCTGACAAATGCTGAGTT 1281
Db 1201 TCTTTCATATAAAGCTTTGAGTTAACTCACTGAGGATATCTGACAAATGCTGAGTT 1260
QY 1282 ATGAACAAAGTGAGGAGAAATGAAATGTAATGCTCTTGAAGAAACATGATGTCGATTT 1341
Db 1261 ATGAACAAAGTGAGGAGAAATGAAATGTAATGCTCTTGAAGAAACATGATGTCGATTT 1320
QY 1342 CAATCCCGTACTTATAAAGAGGTTGGTGAATTTCAAGCTATTTTGGAAATATTTT 1401
Db 1321 CAATCCCGTACTTATAAAGAGGTTGGTGAATTTCAAGCTATTTTGGAAATATTTT 1380
QY 1402 TAGAATATTTTAAAGATTTCAAGCTATTTTCCCTCAAACTCTGAGGAGCTGAGTAACAC 1461
Db 1381 TAGAATATTTTAAAGATTTCAAGCTATTTTCCCTCAAACTCTGAGGAGCTGAGTAACAC 1440
QY 1462 ATCGATCATGATGTAGAGTGTGGTTATGAACCT-----TTAAGTCTTTTATGTTTC 1515
Db 1441 ATCGATCATGATGTAGAGTGTGGTTATGAACCTTTAAGTCTTTTATGTTTCATGTTTC 1500
QY 1516 TATAATAAAGAGTGTCTGC 1536
Db 1501 TATAATAAAGAGTGTCTGC 1521

```

```

RESULT 7
BC013919
LOCUS
DEFINITION Homo sapiens. Similar to thymidylate synthetase, clone MGC:22884
IMAGE:4048625, mRNA, complete cds.
ACCESSION BC013919
VERSION BC013919.1 GI:15530266
KEYWORDS MGC.

```

Sequencing Center (MISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: misc_mgc@nhgri.nih.gov
Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M.,
Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,
Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,
Lim, M., Maduro, O.L., Masiello, C., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Snyder, B., Stantripop, S., Thomas, P.J.,
Tiongson, B.B., Touchman, J.W., Tsurgueon, C., Vogt, J.L., Walker, N.A.,
Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL, Place: 4 Row: P Column: 19
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 4507750.

FEATURES

source

1. 1533
/organism="Homo sapiens"
/db_xref="LocusID:7298"
/db_xref="taxon:9606"
/clone="MGC:1590 IMAGE:3138877"
/tissue type="placenta, choriocarcinoma"
/clone_lib="NIH MGC 21"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
78. 1019
/codon_start=1
/product="thymidylate synthetase"
/protein_id="AAH02567.1"
/db_xref="GI:12803483"
/translation="MPVAGSELPRPLPPAQAQEDAEPRPHGELQYLQIHLRCG
VRKDRGTGSLVFGQARYSLRLEFFLLATKEVFWKGLVLELLFWKINGSTNAKELS
SKGVLDWANGSRDFDLSLGFSTREGLDIPVGFOWHPGAEYRDMSDYSGGVQD
LQVLDITKTNPDRLIMCANPRDLPLMLPCHALCOFVYVNSELSCLYORSQD
MGLVFNFIASVALLFMIAHTGLKPGDFIHTLGDHIVLNHIEPLKIQLOREPRP
PKRLRLKVEKIDDFKAEQFQLEGYNPHPTIRMEMAV"
BASE COUNT 409 a 359 c 389 g 376 t

CDS

Query Match 97.4%; Score 1496.4; DB 9; Length 1533;
Best Local Similarity 99.7%; Pred. No. 5.1e-300;
Matches 1497; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 36 GCGCGCCACTTGGCTGCTCGCTCCGCGCGCCACTTTCGCCCTGCTCCGCCCGC 95
DB 8 GCGCGCCACTTGGCTGCTCGCTCCGCGCGCCACTTTCGCCCTGCTCCGCCCGC 67
QY 96 CCGCGCGCCATGCTGCTGGCGGCTCGAGCTGCGCGCGCGCCCTTTCGCCCGCGC 155
DB 68 CCGCGCGCCATGCTGCTGGCGGCTCGAGCTGCGCGCGCGCCCTTTCGCCCGCGC 127
QY 156 ACAGAGCGGGACCGCGAGCGCGCTCGCGCGCGCGAGCTCGAGTACTTGGGGAGAT 215
DB 128 ACAGAGCGGGACCGCGAGCGCGCTCGCGCGCGCGAGCTGCGAGTACTTGGGGAGAT 187
QY 216 CCACACATCTCCGCTCGCGGCTCAGAGAGGACGACCGCGCGCGCCACCGCCCTGTC 275
DB 188 CCACACATCTCCGCTCGCGGCTCAGAGAGGACGACCGCGCGCGCCACCGCCCTGTC 247
QY 276 GGTATTTCGGCATGCGCGCGCTACGCTGAGAGATGAATTCCTCTGCTGCACACCAA 335
DB 248 GGTATTTCGGCATGCGCGCGCTACGCTGAGAGATGAATTCCTCTGCTGCACACCAA 307
QY 336 ACCTGTGTTCTGGAAGGGTGTGTTGGAGAGTGTGCTGTTATCAAGGATCCACAA 395
DB 308 ACCTGTGTTCTGGAAGGGTGTGTTGGAGAGTGTGCTGTTATCAAGGATCCACAA 367
QY 396 TGCTAAAGAGCTGTCTTCCAGGAGGTGAATCTGGGATGCCAATGGATCCCGAGACTT 455
DB 368 TGCTAAAGAGCTGTCTTCCAGGAGGTGAATCTGGGATGCCAATGGATCCCGAGACTT 427

QY 456 TTTGGACAGCTTGGGATTTCTCCACAGAGAAGAGGGAGCTTGGGCCAGTTTATGGCTT 515
DB 428 TTTGGACAGCTTGGGATTTCTCCACAGAGAAGAGGGAGCTTGGGCCAGTTTATGGCTT 487
QY 516 CCACTGGAGGCAATTTGGGGGAGATACAGAGATATGGAATCAGATTAATTCAGGACAGG 575
DB 488 CCACTGGAGGCAATTTGGGGGAGATACAGAGATATGGAATCAGATTAATTCAGGACAGG 547
QY 576 AGTTGACCAACTGCAAGAGTGAATGACACCATCAACCAACCTTCAGCAGAGAGAT 635
DB 548 AGTTGACCAACTGCAAGAGTGAATGACACCATCAACCAACCTTCAGCAGAGAGAT 607
QY 636 CATCATGTGCGCTTGGAAATCCAGAGATCTTCTCTGATGGCGCTGCTCCATGCCATGC 695
DB 608 CATCATGTGCGCTTGGAAATCCAGAGATCTTCTCTGATGGCGCTGCTCCATGCCATGC 667
QY 696 CTTCTGCCAGTTCTATGTGTGGAACAGTGAAGTGTCTGCGAGCTGTACAGAGATCGGG 755
DB 668 CTTCTGCCAGTTCTATGTGTGGAACAGTGAAGTGTCTGCGAGCTGTACAGAGATCGGG 727
QY 756 AGACATGGCGCTCGGTGTGCGCTTTCAACATCGCAGCTACGCCCTGCTGAGTACATAT 815
DB 728 AGACATGGCGCTCGGTGTGCGCTTTCAACATCGCAGCTACGCCCTGCTGAGTACATAT 787
QY 816 TGGCGCATCATCGCGGCTTGAAGCCAGTGAATTTATACACACTTTGGGAGATGCATAT 875
DB 788 TGGCGCATCATCGCGGCTTGAAGCCAGTGAATTTATACACACTTTGGGAGATGCATAT 847
QY 876 TTACCTGAATCATCATGAGCCACTGAAATTCAGCTTCAGGAGAACCCAGACCTTCC 935
DB 848 TTACCTGAATCATCATGAGCCACTGAAATTCAGCTTCAGGAGAACCCAGACCTTCC 907
QY 936 AAAGCTCAGATTTCTCGAAAGTGAAGAAATTCATGATCTTAAAGTGAAGCTTCA 995
DB 908 AAAGCTCAGATTTCTCGAAAGTGAAGAAATTCATGATCTTAAAGTGAAGCTTCA 967
QY 996 GATTGAGGGTACATCCGATCCCACTATTAAATGAAATGGCTGTTTAGGGTCTTT 1055
DB 968 GATTGAGGGTACATCCGATCCCACTATTAAATGAAATGGCTGTTTAGGGTCTTT 1027
QY 1056 CAAAGAGCTTGAAGGATTTGTGCTTTAGGGGTTGGCTGATGCGAGGTAAAG 1115
DB 1028 CAAAGAGCTTGAAGGATTTGTGCTTTAGGGGTTGGCTGATGCGAGGTAAAG 1087
QY 1116 TTCTTTTGTCTTAAAGAAAGAAAGCACTAGGTCAAAAATCTGTCGGTGAACCTTACGTT 1175
DB 1088 TTCTTTTGTCTTAAAGAAAGAAAGCACTAGGTCAAAAATCTGTCGGTGAACCTTACGTT 1147
QY 1176 ATTAATTTTAAAGATTTGCTCACTGGCAAAATGTAACCTGTGCCAGTTCTTTCCATAATA 1235
DB 1148 ATTAATTTTAAAGATTTGCTCACTGGCAAAATGTAACCTGTGCCAGTTCTTTCCATAATA 1207
QY 1236 AAGCTTTGAGTTAACTCACTGAGGATTCGACATGCTGAGGTATGACAAAGTGAG 1295
DB 1208 AAGCTTTGAGTTAACTCACTGAGGATTCGACATGCTGAGGTATGACAAAGTGAG 1267
QY 1296 GAGAAATGAATGTATGCTCTTAGCAAAAAGATGTATGCTCAATCCCACTACT 1355
DB 1268 GAGAAATGAATGTATGCTCTTAGCAAAAAGATGTATGCTCAATCCCACTACT 1327
QY 1356 TATAAAGAAAGGTTGCTGAATTTCAAGCTATTTTGGAAATTTTGGAAATTTTAAAG 1415
DB 1328 TATAAAGAAAGGTTGCTGAATTTCAAGCTATTTTGGAAATTTTGGAAATTTTAAAG 1387
QY 1416 AATTTCAGAGCTATTCCTCAATCTGAGGAGCTGAGTAACCACTGCATCATGATGT 1475
DB 1388 AATTTCAGAGCTATTCCTCAATCTGAGGAGCTGAGTAACCACTGCATCATGATGT 1447
QY 1476 AGAGTGTGTTATGAACCTTANAGTGTGTTTATATGTTGCTTATATAAAGAGTGTCTG 1535
DB 1448 AGAGTGTGTTATGAACCTTATAGTGTGTTTATATGTTGCTTATATAAAGAGTGTCTG 1507
QY 1536 C 1536

MLGVFPNLIASVALLTYMIAHTIGLKEGDFIHTLGDHAIYILNHIEPLKIQLOREPRFF
PKLRIIRKVEKIDDKAEDFOEGVNPHTIKMEAV"
1231..1236
/note="pot. polyadenylation signal"
1519..1524
/note="pot. polyadenylation signal"
1536
polyA_site
/note="polyadenylation site"
BASE COUNT 390 a 369 c 399 g 378 t
ORIGIN

Query Match 99.8%; Score 1533; DB 9; Length 1536;
Best Local Similarity 99.8%; Pred. No. 1.3e-307;
Matches 1533; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGACACATTGGCTCGCTCCGTCGGCCGGCCGACATGGCGCTCCCTCCGT 60
Db 1 GGGGGGGGGGACACATTGGCTCGCTCCGTCGGCCGGCCGACATGGCGCTCCCTCCGT 60

Qy 61 CCGCCGGGCGCACTTCGGCTCGCTCCGTCGGCCGGCCGCGCCGATGCGCTGGCGCGC 120
Db 61 CCGCCGGGCGCACTTCGGCTCGCTCCGTCGGCCGGCCGCGCCGATGCGCTGGCGCGC 120

Qy 121 TCGAGCTGCGCGCGCGCGCTTCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
Db 121 TCGAGCTGCGCGCGCGCGCTTCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180

Qy 181 CCGCCGCGAGGAGCTGAGTACCTGGGGGAGATCGAACHATCTCCGCTCGCGCGCTC 240
Db 181 CCGCCGCGAGGAGCTGAGTACCTGGGGGAGATCGAACHATCTCCGCTCGCGCGCTC 240

Qy 241 AGAAGAGCAGCGCAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTAC 300
Db 241 AGAAGAGCAGCGCAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTAC 300

Qy 301 AGCTCAGAGATGAAATCCCTCTGCTGACAAACAAAGCTGTTCTGAAAGGCTGTTTG 360
Db 301 AGCTCAGAGATGAAATCCCTCTGCTGACAAACAAAGCTGTTCTGAAAGGCTGTTTG 360

Qy 361 GAGGAGTCTCTGCTGTTTATCAAGGATCCAAATGCTTAAAGAGCTGTCTTCAAGGGA 420
Db 361 GAGGAGTCTCTGCTGTTTATCAAGGATCCAAATGCTTAAAGAGCTGTCTTCAAGGGA 420

Qy 421 GTCAAAATCTGGATGCGCAATGATCCGAGACTTTTGGAGCGCTGGGATTCACAC 480
Db 421 GTCAAAATCTGGATGCGCAATGATCCGAGACTTTTGGAGCGCTGGGATTCACAC 480

Qy 481 AGAGAAGAGGGGACTTGGGCCCGCTTATGCTTCCAGTGAGGCAATTTGGGGCAGAA 540
Db 481 AGAGAAGAGGGGACTTGGGCCCGCTTATGCTTCCAGTGAGGCAATTTGGGGCAGAA 540

Qy 541 TACAGAGATATGGAATCAGATTATCAGGACAGGAGGTTGACCACTGCAAGAGTGATT 600
Db 541 TACAGAGATATGGAATCAGATTATCAGGACAGGAGGTTGACCACTGCAAGAGTGATT 600

Qy 601 GACACATCAAAACCAACCTCGACAGAGAAATCATCTGCTGCTGGATTCACAGA 660
Db 601 GACACATCAAAACCAACCTCGACAGAGAAATCATCTGCTGCTGGATTCACAGA 660

Qy 661 GATCTCTCTGATGGCGCTCCATGCTCCATGCTCCCTCTGCTGCTGCTGCTGCTGCTG 720
Db 661 GATCTCTCTGATGGCGCTCCATGCTCCATGCTCCCTCTGCTGCTGCTGCTGCTGCTG 720

Qy 721 AGTGAGCTGTCTGCGAGCTGTACAGAGATCGGAGACATGGCGCTCGGCTGCTTTC 780
Db 721 AGTGAGCTGTCTGCGAGCTGTACAGAGATCGGAGACATGGCGCTCGGCTGCTTTC 780

Qy 781 AACATCGCAGCTACGCTCTCTCAGTATGCTGATGCTGCTGCTGCTGCTGCTGCTGCT 840
Db 781 AACATCGCAGCTACGCTCTCTCAGTATGCTGATGCTGCTGCTGCTGCTGCTGCTGCT 840

Qy 841 GGTGACTTTATACACATTTGGGAGATGACATATTTACTGAAATCATCGACCCACTG 900

Db 841 GGTGACTTTATACACATTTGGGAGATGACATATTTACTTGAATCATCGACCCACTG 900

Qy 901 AAAATTCAGCTTCAGCGAGAACCCAGAGCTTTTCCAAAGCTCAGGATTTCTTGGAAAAGTT 960

Db 901 AAAATTCAGCTTCAGCGAGAACCCAGAGCTTTTCCAAAGCTCAGGATTTCTTGGAAAAGTT 960

Qy 961 GAGAAAATTCAGTCACTTCAAAAGCTGGAAGACTTTTCAAGTTGAAGGCTCAATTCGCCATCCA 1020

Db 961 GAGAAAATTCAGTCACTTCAAAAGCTGGAAGACTTTTCAAGTTGAAGGCTCAATTCGCCATCCA 1020

Qy 1021 ACTATTAAAAATGAAAAATGGCTTTTAGGGTCTTCCAAAGGAGCTGGAAGGATTTGCTCA 1080

Db 1021 ACTATTAAAAATGAAAAATGGCTTTTAGGGTCTTCCAAAGGAGCTTGAAGGATTTGCTCA 1080

Qy 1081 GTCTTTAGGGGTTGGGCTGGATGCGGAGGTAAAGTTCTTTTGGCTCTTAAAGAAAGG 1140

Db 1081 GTCTTTAGGGGTTGGGCTGGATGCGGAGGTAAAGTTCTTTTGGCTCTTAAAGAAAGG 1140

Qy 1141 AACTAGTCAAAAAATCTGTCGTCGACCTATCAGTTATTAATTTTAAAGGATTTGCCACT 1200

Db 1141 AACTAGTCAAAAAATCTGTCGTCGACCTATCAGTTATTAATTTTAAAGGATTTGCCACT 1200

Qy 1201 GGCATAATGTAATGTCGCGAGTCTTCCATAATAAAGGCTTTGAGTTAACTCACTGAGG 1260

Db 1201 GGCATAATGTAATGTCGCGAGTCTTCCATAATAAAGGCTTTGAGTTAACTCACTGAGG 1260

Qy 1261 GTATCTGCAATGCTGAGGTATGACAAAGTGAGGAGATGAAATGATGTGCTCTTAG 1320

Db 1261 GTATCTGCAATGCTGAGGTATGACAAAGTGAGGAGATGAAATGATGTGCTCTTAG 1320

Qy 1321 CAAAAACATGATGTGCTCAATTTCAATCCACGTTACTTATAAAGAGGTTGGTGAATTTAC 1380

Db 1321 CAAAAACATGATGTGCTCAATTTCAATCCACGTTACTTATAAAGAGGTTGGTGAATTTAC 1380

Qy 1381 AAGCTATTTTGGAAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAAT 1440

Db 1381 AAGCTATTTTGGAAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAAT 1440

Qy 1441 CTGAGGAGCTGAGTAAACCATCGATCATGATGATGATGATGATGATGATGATGATGATGAT 1500

Db 1441 CTGAGGAGCTGAGTAAACCATCGATCATGATGATGATGATGATGATGATGATGATGATGAT 1500

Qy 1501 TGTATTATATGTTGCTATATAAAGAGGTTCTTCG 1536

Db 1501 TGTATTATATGTTGCTATATAAAGAGGTTCTTCG 1536

RESULT 5

BC002567

LOCUS

DEFINITION

ACCSSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REMARK

COMMENT

BC002567 1533 bp mRNA linear PRI 12-JUL-2001
Homo sapiens, thymidylate synthetase, clone MGC:1590 IMAGE:3138877,
mRNA, complete cds.
BC002567 GI:12803482
BC002567.1
MGC.
Homo sapiens.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 1533)
Strausberg, R.
Direct Submission
Submitted (05-FEB-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cdNA Library Preparation: Rubin Laboratory
cdNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural

Qy	241	AGGAAGGACGACCGACCGGACCGGACCGCTGTGGTATTCGGCATGAGGCGCGCTAC	300
Db	241	AGGAAGGACGACCGACCGGACCGGACCGCTGTGGTATTCGGCATGAGGCGCGCTAC	300
Qy	301	AGCCTGAGAGATGAATTCCTCTCTGCTGACAAACAAACGTGTGTTCTGGAAGGTTGTTTG	360
Db	301	AGCCTGAGAGATGAATTCCTCTCTGCTGACAAACAAACGTGTGTTCTGGAAGGTTGTTTG	360
Qy	361	GAGGAGTTGCTGTGTTTATCAAGGATCCAAATGCTAAAGAGCTGTCTTCCAAAGGGA	420
Db	361	GAGGAGTTGCTGTGTTTATCAAGGATCCAAATGCTAAAGAGCTGTCTTCCAAAGGGA	420
Qy	421	GTGAAAATCTGGGATGCCAATGGATCCGAGACTTTTGGACAGCCTGGGATTCCTCACC	480
Db	421	GTGAAAATCTGGGATGCCAATGGATCCGAGACTTTTGGACAGCCTGGGATTCCTCACC	480
Qy	481	AGAGAAGAGGGGACTTGGGCCCGAGTTATAGCTTCAGTGGAGGACATTTGGGCGACAA	540
Db	481	AGAGAAGAGGGGACTTGGGCCCGAGTTATAGCTTCAGTGGAGGACATTTGGGCGACAA	540
Qy	541	TACAGAGATATGGAATCAGATTATTACGACAGGGAGTTGACCACTGCAAAAGAGTGATT	600
Db	541	TACAGAGATATGGAATCAGATTATTACGACAGGGAGTTGACCACTGCAAAAGAGTGATT	600
Qy	601	GACACCATCAAAAACCAACCCCTGACGACAGAGAATCATCATGTGCGCTTGGAAATCCAAAG	660
Db	601	GACACCATCAAAAACCAACCCCTGACGACAGAGAATCATCATGTGCGCTTGGAAATCCAAAG	660
Qy	661	GATCTTCCTCTCATGGCGCTGCCTCCATGCCATGCCCTCTGCGAGTTCTATGTGTGAAC	720
Db	661	GATCTTCCTCTCATGGCGCTGCCTCCATGCCATGCCCTCTGCGAGTTCTATGTGTGAAC	720
Qy	721	AGTGAGCTGTCTGCCAGCTGTACAGAGATCGGAGACATGCGGCTCGGTGTGCTTTTC	780
Db	721	AGTGAGCTGTCTGCCAGCTGTACAGAGATCGGAGACATGCGGCTCGGTGTGCTTTTC	780
Qy	781	AACATCGCAGCTACGCCCTGCTCACTGATGATGCGACATCAGCGGCTGTGAGCCA	840
Db	781	AACATCGCAGCTACGCCCTGCTCACTGATGATGCGACATCAGCGGCTGTGAGCCA	840
Qy	841	GGTGACTTTATACACATCTTTGGGAGTGACATATTTACCTGAAATCACATCGAGCCACTG	900
Db	841	GGTGACTTTATACACATCTTTGGGAGTGACATATTTACCTGAAATCACATCGAGCCACTG	900
Qy	901	AAAAATTCAGCTTCAGGAGAACCCAGACCTTTCCAAAGCTCAGGATCTTCGAAAAGTT	960
Db	901	AAAAATTCAGCTTCAGGAGAACCCAGACCTTTCCAAAGCTCAGGATCTTCGAAAAGTT	960
Qy	961	GAGAAAATGATGACTTCAAAGCTGGAAGACTTTTCAGATTGAAGGTTACAATCCGATCCA	1020
Db	961	GAGAAAATGATGACTTCAAAGCTGGAAGACTTTTCAGATTGAAGGTTACAATCCGATCCA	1020
Qy	1021	ACTATTAAATGGAATGGCTGTTTAGGTGCTTTCAAAGGAGCTGGAAGGATATTGTCA	1080
Db	1021	ACTATTAAATGGAATGGCTGTTTAGGTGCTTTCAAAGGAGCTTGAAGGATATTGTCA	1080
Qy	1081	GTCTTTAGGGTTGGGCTGGATCCGAGTTAAAGCTTTCTTTTGTCTCTAAAGANAAGG	1140
Db	1081	GTCTTTAGGGTTGGGCTGGATCCGAGTTAAAGCTTTCTTTTGTCTCTAAAGANAAGG	1140
Qy	1141	AACTAGGTCAAAAATCTGTCCGTGACTATCAGTTATTAATTTTAAAGGATGTGCGACT	1200
Db	1141	AACTAGGTCAAAAATCTGTCCGTGACTATCAGTTATTAATTTTAAAGGATGTGCGACT	1200
Qy	1201	GGCAAATGTAACTGTGCCAGTTCTTTCCATATAAAGGCTTTGAGTTAACTCACTGAGG	1260
Db	1201	GGCAAATGTAACTGTGCCAGTTCTTTCCATATAAAGGCTTTGAGTTAACTCACTGAGG	1260
Qy	1261	GTATCTGACATGCTGAGGTTATGAACAAAGTGAGGAGATGAAATGTATGTCTCTTAG	1320
Db	1261	GTATCTGACATGCTGAGGTTATGAACAAAGTGAGGAGATGAAATGTATGTCTCTTAG	1320

[illegible]

Sun Dec 1 18:15:45 2002

/tissue type="Bone marrow, chronic myelogenous leukemia"
/clone lib="NIH MGC_54"
/lab_host="DH10B"
/note="Vector: pDNR-LIB"
96...1037
/codon_start=1
/product="similar to thymidylate synthetase"
/protein_id="AAH13919.1"
/db_xref="GI:15530267"
/translation="MPVAGSELPRP...PPAAQERDAERPPHGLQVLQIQLRLTG
VRKDDTGTLSVFCMGARSLRDEPLLTKRVFKVLELLNFIKSTNAKES
SKGKIDANGSDFLDLGFSTREBGLGVYQFQWRFGBAYRDMESYSQGVQ
LQVIDITIKNPDRRIIMCAWNRDLPLMALPPCHALCOFVNVNSLSQLYQRSGD
MGLVGFPIASVALLTYMIAHTGLKPGDFIHLGLDAHYLNHLEPLKIQLOREPRP
PKRLILRKVEKIDDFKAEQDFQEGYNPHPTIKEMAV"
BASE COUNT 425 a 368 c 398 g 378 t

Query Match	88.4%	Score 1358	DB 9	Length 1569	0
Best Local Similarity	99.9%	Pred. No. 0			
Matches 1458	Conservative	0	Mismatches	2	Indels
					Gaps
QY	37	CCGCGCCACTTGGCCTCGCTCCGCGCGCCACTTGCCTGCGCTCCGCTCCCGCCG	96		
Db	27	CCGCGCCACTTGGCCTCGCTCCGCGCGCCACTTGCCTGCGCTCCGCTCCCGCCG	86		
QY	97	CGCGCGCCATGCGCTCGGAGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	156		
Db	87	CGCGCGCCATGCGCTCGGAGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	146		
QY	157	CAGGAGCGGAGCG	216		
Db	147	CAGGAGCGGAGCG	206		
QY	217	CAACACATCTCGCTCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	276		
Db	207	CAACACATCTCGCTCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	266		
QY	277	GTATTCGGGATGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	336		
Db	267	GTATTCGGGATGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	326		
QY	337	CGTGTCTTGGAGGCTGTTTGGAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCT	396		
Db	327	CGTGTCTTGGAGGCTGTTTGGAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCT	386		
QY	397	GCTAAAGAGCTGTCTTCCAGGAGGTGAAATCTGGGATGCAATGATGATGATGAT	456		
Db	387	GCTAAAGAGCTGTCTTCCAGGAGGTGAAATCTGGGATGCAATGATGATGATGAT	446		
QY	457	TTGGACAGCTGGGATCTCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	516		
Db	447	TTGGACAGCTGGGATCTCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	506		
QY	517	CAGTGGAGGCAATTTGGGGGAGAGATACAGAGATATGGAATGGAATGGAATGGA	576		
Db	507	CAGTGGAGGCAATTTGGGGGAGAGATACAGAGATATGGAATGGAATGGAATGGA	566		
QY	577	GTTGACCAACTGCAAGAGTATTTGACCAATCAAAACCAACCTGACGACAGAGATC	636		
Db	567	GTTGACCAACTGCAAGAGTATTTGACCAATCAAAACCAACCTGACGACAGAGATC	626		
QY	637	ATCATGTGGCTTGAATTCGAGAGATCTTCTGATGCGCTGCTGCTGCTGCTGCTG	696		
Db	627	ATCATGTGGCTTGAATTCGAGAGATCTTCTGATGCGCTGCTGCTGCTGCTGCTG	686		
QY	697	CTCTGCGAGTCTATGTGTGAACAGTACGCTGTCTGCGAGCTGTCTGCGAGCTGT	756		
Db	687	CTCTGCGAGTCTATGTGTGAACAGTACGCTGTCTGCGAGCTGTCTGCGAGCTGT	746		
QY	757	GACATGGGCTCGGTGTGCTTTTCAACATTCGCGAGCTGTCTGCGAGCTGTCTGCG	816		
Db	747	GACATGGGCTCGGTGTGCTTTTCAACATTCGCGAGCTGTCTGCGAGCTGTCTGCG	806		

QY	817	GGGCACATCAACGGGCTGAGCCAGGTGACTTTATACACATTTGGGAGTGCATATT	876
Db	807	GGGCACATCAACGGGCTGAGCCAGGTGACTTTATACACATTTGGGAGTGCATATT	866
QY	877	TACCTGAATCAATCGAGGCTGAAATTCAGCTTCAGGAGACCCAGACCTTCCCA	936
Db	867	TACCTGAATCAATCGAGGCTGAAATTCAGCTTCAGGAGACCCAGACCTTCCCA	926
QY	937	TACCTGAATCAATCGAGGCTGAAATTCAGCTTCAGGAGACCCAGACCTTCCCA	996
Db	927	TACCTGAATCAATCGAGGCTGAAATTCAGCTTCAGGAGACCCAGACCTTCCCA	986
QY	997	ATTGAAGGTACAAATCGGATCCCACTTAAATAGGAATGGCTGTGGTGGTTC	1056
Db	987	ATTGAAGGTACAAATCGGATCCCACTTAAATAGGAATGGCTGTGGTGGTTC	1046
QY	1057	AAAGGAGCTGAGGATATTTGCTAGCTCTTTAGGGTTGGGCTGGATGCGGATTAAGT	1116
Db	1047	AAAGGAGCTGAGGATATTTGCTAGCTCTTTAGGGTTGGGCTGGATGCGGATTAAGT	1106
QY	1117	TCTTTTCTCTTAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGT	1176
Db	1107	TCTTTTCTCTTAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGT	1166
QY	1177	TCTTTTCTCTTAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGT	1236
Db	1167	TCTTTTCTCTTAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGT	1226
QY	1237	AGGCTTGGAGTTAACTCACTGAGGCTGATCAATGCTGAGGTTATGACAAAGTGGG	1296
Db	1227	AGGCTTGGAGTTAACTCACTGAGGCTGATCAATGCTGAGGTTATGACAAAGTGGG	1286
QY	1297	AGGCTTGGAGTTAACTCACTGAGGCTGATCAATGCTGAGGTTATGACAAAGTGGG	1356
Db	1287	AGGCTTGGAGTTAACTCACTGAGGCTGATCAATGCTGAGGTTATGACAAAGTGGG	1346
QY	1357	ATTAAGAGGCTGGGTAATTTCAACAGCTATTTTGGATATTTTGAATATTTTAAGA	1416
Db	1347	ATTAAGAGGCTGGGTAATTTCAACAGCTATTTTGGATATTTTGAATATTTTAAGA	1406
QY	1417	ATTTTCAACAGCTATTTTCCCTCAAAATCTGAGGAGCTGAGTAAACATGATGATTA	1476
Db	1407	ATTTTCAACAGCTATTTTCCCTCAAAATCTGAGGAGCTGAGTAAACATGATGATTA	1466
QY	1477	GAGTGTGGTATGAACTTTA 1496	
Db	1467	GAGTGTGGTATGAACTTTA 1486	

RESULT 6
BC002567 1533 bp mRNA linear PRI 12-JUL-2002
LOCUS Homo sapiens, thymidylate synthetase, clone MGC:1590 IMAGE:31388
DEFINITION mRNA, complete cds.
ACCESSION BC002567
VERSION BC002567.1 GI:12803482
KEYWORDS MGC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1533)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (05-FEB-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
REMARK COMMENT

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

COM nucleic - nucleic search, using sw model

Run on: November 28, 2002, 19:48:24 ; Search time 119.586 Seconds
(without alignments)
8693.659 Million cell updates/sec

Title: US-09-963-333-6
Perfect score: 1536
Sequence: 1 gggggggggggggaccacttg.....ataataaagaagtgttctgc 1536

Scoring table: IDENTITY_NUC
Gapop 10.0 Gapext 1.0

Searched: 344316 seqs, 338423730 residues

total number of hits satisfying chosen parameters: 688632

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing:	Minimum Match	0%
	Maximum Match	100%
	Listing first	45 s

Database : Pending Patents NA New: *

[illegible]

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is generated by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1533	99.8	1536	5	US-09-658-6598-6
2	1493.4	97.2	3298	6	US-10-240-965-200
3	1126.4	73.3	1942	5	US-09-724-676-34523
C 4	1126.4	73.3	1942	5	US-09-724-676-34523
C 5	1126.4	73.3	2042	5	US-09-724-676-34515
C 6	1126.4	73.3	2042	5	US-09-724-676A-34515
C 7	1126.4	73.3	2391	5	US-09-724-676-34597
C 8	1126.4	73.3	2391	5	US-09-724-676A-34597
C 9	1126.4	73.3	2411	5	US-09-724-676-34535
C 10	1126.4	73.3	2411	5	US-09-724-676A-34535
C 11	1126.4	73.3	2491	5	US-09-724-676-34589
C 12	1126.4	73.3	2491	5	US-09-724-676A-34589
C 13	1126.4	73.3	2527	5	US-09-724-676-34548
C 14	1126.4	73.3	2527	5	US-09-724-676A-34548
C 15	1126.4	73.3	3023	5	US-09-724-676-34565
C 16	1126.4	73.3	3023	5	US-09-724-676A-34565
C 17	1126.4	73.3	3123	5	US-09-724-676-34557
C 18	1126.4	73.3	3123	5	US-09-724-676A-34557
C 19	1126.4	73.3	3139	5	US-09-724-676-34581
C 20	1126.4	73.3	3139	5	US-09-724-676A-34581
C 21	1126.4	73.3	3239	5	US-09-724-676-34573
C 22	1126.4	73.3	3239	5	US-09-724-676A-34573
C 23	928.8	60.5	2083	5	US-09-724-676-34524
C 24	928.8	60.5	2083	5	US-09-724-676A-34524
C 25	928.8	60.5	2183	5	US-09-724-676-34516
C 26	938.9	60.5	2183	5	US-09-724-676A-34516

C 27	928.8	60.5	2552	5	US-09-T24-676-34536	Sequence 34536, A
C 28	928.8	60.5	2552	5	US-09-T24-676A-34536	Sequence 34536, A
C 29	928.8	60.5	2552	5	US-09-T24-676A-34590	Sequence 34590, A
C 30	928.8	60.5	2632	5	US-09-T24-676A-34590	Sequence 34590, A
C 31	928.8	60.5	2652	5	US-09-T24-676-34529	Sequence 34529, A
C 32	928.8	60.5	2652	5	US-09-T24-676A-34529	Sequence 34529, A
C 33	928.8	60.5	2668	5	US-09-T24-676-34549	Sequence 34549, A
C 34	928.8	60.5	2668	5	US-09-T24-676A-34549	Sequence 34549, A
C 35	928.8	60.5	2768	5	US-09-T24-676-34542	Sequence 34542, A
C 36	928.8	60.5	2768	5	US-09-T24-676A-34542	Sequence 34542, A
C 37	928.8	60.5	3164	5	US-09-T24-676-34566	Sequence 34566, A
C 38	928.8	60.5	3164	5	US-09-T24-676A-34566	Sequence 34566, A
C 39	928.8	60.5	3253	5	US-09-T24-676-34525	Sequence 34525, A
C 40	928.8	60.5	3253	5	US-09-T24-676A-34525	Sequence 34525, A
C 41	928.8	60.5	3264	5	US-09-T24-676-34558	Sequence 34558, A
C 42	928.8	60.5	3264	5	US-09-T24-676A-34558	Sequence 34558, A
C 43	928.8	60.5	3280	5	US-09-T24-676-34582	Sequence 34582, A
C 44	928.8	60.5	3280	5	US-09-T24-676A-34582	Sequence 34582, A
C 45	928.8	60.5	3353	5	US-09-T24-676-34517	Sequence 34517, A
C 46	928.8	60.5	3353	5	US-09-T24-676A-34517	Sequence 34517, A

ATTACHMENTS

RESULT 1

```

US-09-658-659B-6
; Sequence 6, Application US/09658659B
; GENERAL INFORMATION:
; APPLICANT: Stanton, Jr., Vincent P.
; TITLE OF INVENTION: GENE SEQUENCE VARIANCE IN GENES RELATED
; TO FOLATE METABOLISM HAVING UTILITY IN DETERMINING THE
; TITLE OF INVENTION: TREATMENT OF DISEASE
; TITLE OF INVENTION: TREATMENT OF DISEASE
; FILE REFERENCE: 11926-015001
; CURRENT APPLICATION NUMBER: US/09/658,659B
; CURRENT FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 09/596,033
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 09/357,743
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 09/357,024
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: 60/093,484
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1536

```

LENGTH: 1336
DATE: 11/11/58

```

1 TYPE: DNA
2 ORGANISM: Homo sapiens
3
4 FEATURE:
5     NAME/KEY: misc_feature
6     LOCATION: 1066
7     OTHER INFORMATION: n = t o r c
8
9     FEATURE:
10        NAME/KEY: misc_feature
11        LOCATION: 1136
12        OTHER INFORMATION: n = a o r g
13
14        FEATURE:
15            NAME/KEY: misc_feature
16            LOCATION: 1497
17            OTHER INFORMATION: n = t o r a
18
19 US-09-6558-6598-6

```

Source: Match

qq	8%	score	1533:	DB	5:	Length	1536;
----	----	-------	-------	----	----	--------	-------

Query Match	99.8%;
Best Local Similarity	100.0%;
Mismatches	0;
Indels	0;
Gaps	0;

1 GGGGGGGGGGACCACTTGGCCCTGCCCTCCGTCGGCGGCCACTTGGCCCTGCCCTCCGT 60

1 GGGGGGGGACCACTTGGCCTGCCCTCCGTCCGCGCGCCACTTGGCCTGCCCTCCGT 60

CGGCGCCATGCCCTGTGGCCGC 120

QY
T9
T9


```
Db 1522 GTCAAAATCTGGGATGCCAATSGATCCGAGACTTTTGGACAGCCTGGGATTTCTCACC 1463
QY 481 AGAGAGAGAGGAGACTTGGGCCCGAGTTTATGGCTTCCAGTGGAGCAATTTGGGGAGAA 540
Db 1462 AGAGAGAGAGGAGACTTGGGCCCGAGTTTATGGCTTCCAGTGGAGCAATTTGGGGAGAA 1403
QY 541 TACAGAGATATGAATCAGATTATTCAGGACAGGAGTTCACCACTGCAAAAGTGATT 600
Db 1402 TACAGAGATATGAATCAGATTATTCAGGACAGGAGTTCACCACTGCAAAAGTGATT 1343
QY 601 GACACCATCAAAACCAACCCCTGACGACAGAGAAATCATCATGTGCGCTTGGATTCGAAGA 660
Db 1342 GACACCATCAAAACCAACCCCTGACGACAGAGAAATCATCATGTGCGCTTGGATTCGAAGA 1283
QY 661 GATCTTCTCTGATGGGCGCTTCGATGCGATGCGCTTCCAGCTTCTATGTGGTGAAC 720
Db 1282 GATCTTCTCTGATGGGCGCTTCGATGCGATGCGCTTCCAGCTTCTATGTGGTGAAC 1223
QY 721 AGTGAGCTGTCTGACCTGTACGAGATCGGAGACATCGGCGCTCGGTGCTTTTC 780
Db 1222 AGTGAGCTGTCTGACCTGTACGAGATCGGAGACATCGGCGCTCGGTGCTTTTC 1163
QY 781 AACATCGCCAGCTAGCCCTCTCACTGATCATGATTGCGCATCATCGGCGCTGAAGCCA 840
Db 1162 AACATCGCCAGCTAGCCCTCTCACTGATCATGATTGCGCATCATCGGCGCTGAAGCCA 1103
QY 841 GGTGACTTTATACACATTTGGGAGATGACATATTTACTGAAATCATCATCGAGCACTG 900
Db 1102 GGTGACTTTATACACATTTGGGAGATGACATATTTACTGAAATCATCATCGAGCACTG 1043
QY 901 AAAATTCACTTTCAGGAGAACCCAGACCTTTCCAAAGCTCAGGATTTTCGAAAAGTT 960
Db 1042 AAAATTCACTTTCAGGAGAACCCAGACCTTTCCAAAGCTCAGGATTTTCGAAAAGTT 983
QY 961 GAGAAAATGATGACTTCAAGACTCAAGACTTTTCCAAAGCTCAGGATTTTCGAAAAGTT 1020
Db 982 GAGAAAATGATGACTTCAAGACTCAAGACTTTTCCAAAGCTCAGGATTTTCGAAAAGTT 923
QY 1021 ACTATTAATGAAATGGCTTTTGGGTGCTTTCAAGAGCTGAGGATTTTCGAAAAGTT 1080
Db 922 ACTATTAATGAAATGGCTTTTGGGTGCTTTTCAAGAGCTGAGGATTTTCGAAAAGTT 863
QY 1081 GTCCTTAGGGGTGGGCTGGATGCGGAGTAAAGTTCTTTTGGCTCTTAAAGANAAGG 1140
QY 1141 AACTAGTCAAAAATCTCTCCG 1162
Db 802 AGGGTGTATGACTTCTAGCTG 781

RESULT 4
US-09-724-676A-34523/c
; Sequence 34523, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Comugen Ltd
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Comugen
; CURRENT APPLICATION NUMBER: US/09724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 34523
; LENGTH: 1942
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676A-34523

Query Match 73.3%; Score 1126.4; DB 5; Length 1942;
Best Local Similarity 98.0%; Pred. No. 2.6e-275;
Matches 1139; Conservative 0; Mismatches 23; Indels 0; Gaps 0;
QY 1 GGGGGGGGGGAGCACTTGGCTGCTCCGTCGGCGGCCCACTTGGCTGCTCCGT 60
```

```
Db 1942 GGGGGGGGGGAGCACTTGGCTGCTCCGTCGGCGGCCCACTTGGCTGCTCCGT 1883
QY 61 CCGCGCCGCGCACTTTGGCTGCTCCGTCGGCGGCCCGCGGCCCATGCTGTGGCGCGC 120
Db 1882 CCGCGCCGCGCACTTTGGCTGCTCCGTCGGCGGCCCGCGGCCCATGCTGTGGCGCGC 1823
QY 121 TCGGAGTGTCCGCGCGCCCTTCCCTCCCGCGCACAGGAGCGCGGACGCGCGCGCT 180
Db 1822 TCGGAGTGTCCGCGCGCCCTTCCCTCCCGCGCACAGGAGCGCGGACGCGCGCGCT 1763
QY 181 CCGCGCGCACGCGGAGTGTGAGTACTTGGGCGAGATCCACACATCTCCGCTGGCGGCTC 240
Db 1762 CCGCGCGCACGCGGAGTGTGAGTACTTGGGCGAGATCCACACATCTCCGCTGGCGGCTC 1703
QY 241 AGGAGGACGACGCGGCGCACCGGCGCACCGCTGCTGGTATTCGGCATGCGGCGCGCTAC 300
Db 1702 AGGAGGACGACGCGGCGCACCGGCGCACCGCTGCTGGTATTCGGCATGCGGCGCGCTAC 1643
QY 301 AGCTGAGAGATGAATTCCTCTCTGACAAACCAACGCTGTCTTCTGGAAGGCTGTTTG 360
Db 1642 AGCTGAGAGATGAATTCCTCTCTGACAAACCAACGCTGTCTTCTGGAAGGCTGTTTG 1583
QY 361 GAGGAGTGTCTGTGTTTATCAAGGAGTCCAAATGCTTAAAGAGCTGTCTTCCAGGGA 420
Db 1582 GAGGAGTGTCTGTGTTTATCAAGGAGTCCAAATGCTTAAAGAGCTGTCTTCCAGGGA 1523
QY 421 GTGAAAATCTGGGATGCGAATGAGATCCGAGACTTTTGGACAGCCTGGGATTTCTCACC 480
Db 1522 GTGAAAATCTGGGATGCGAATGAGATCCGAGACTTTTGGACAGCCTGGGATTTCTCACC 1463
QY 481 AGAGAGAGGAGGACTTGGGCGCGAGTTTATGGCTTCCAGTGGAGCAATTTGGGGCAGAA 540
Db 1462 AGAGAGAGGAGGACTTGGGCGCGAGTTTATGGCTTCCAGTGGAGCAATTTGGGGCAGAA 1403
QY 541 TACAGAGATATGAATCAGATTATTCAGGACAGGAGTTCACCACTGCAAAAGTGATT 600
Db 1402 TACAGAGATATGAATCAGATTATTCAGGACAGGAGTTCACCACTGCAAAAGTGATT 1343
QY 601 GACACCATCAAAACCAACCCCTGACGACAGAGAAATCATCATGTGCGCTTGGAAATCCAAGA 660
Db 1342 GACACCATCAAAACCAACCCCTGACGACAGAGAAATCATCATGTGCGCTTGGAAATCCAAGA 1283
QY 661 GATCTTCTCTGATGGCGCTGCTCCATGCCATCCCTCTGCGAGTTCATGTGCTGAAC 720
Db 1282 GATCTTCTCTGATGGCGCTGCTCCATGCCATCCCTCTGCGAGTTCATGTGCTGAAC 1223
QY 721 AGTGAGTGTCTGCTCCAGCTGTACAGAGATCGGAGACATGGGCGCTCGGTGCTTTTC 780
Db 1222 AGTGAGTGTCTGCTCCAGCTGTACAGAGATCGGAGACATGGGCGCTCGGTGCTTTTC 1163
QY 781 AACATCGCCAGCTAGCCCTCTCACTGATCATGATTGCGCATCATCGGCGCTGAAGCCA 840
Db 1162 AACATCGCCAGCTAGCCCTCTCACTGATCATGATTGCGCATCATCGGCGCTGAAGCCA 1103
QY 841 GGTGACTTTATACACATTTGGGAGATGACATATTTACTGAAATCATCATCGAGCACTG 900
Db 1102 GGTGACTTTATACACATTTGGGAGATGACATATTTACTGAAATCATCATCGAGCACTG 1043
QY 901 AAAATTCACTTTCAGGAGAACCCAGACCTTTCCAAAGCTCAGGATTTTCGAAAAGTT 960
Db 1042 AAAATTCACTTTCAGGAGAACCCAGACCTTTCCAAAGCTCAGGATTTTCGAAAAGTT 983
QY 961 GAGAAAATGATGACTTCAAGACTCAAGACTTTTCCAAAGCTCAGGATTTTCGAAAAGTT 1020
Db 982 GAGAAAATGATGACTTCAAGACTCAAGACTTTTCCAAAGCTCAGGATTTTCGAAAAGTT 923
QY 1021 ACTATTAATGAAATGGCTTTTGGGTGCTTTTCAAGAGCTGAGGATTTTCGAAAAGTT 1080
Db 922 ACTATTAATGAAATGGCTTTTGGGTGCTTTTCAAGAGCTGAGGATTTTCGAAAAGTT 863
QY 1081 GTCCTTAGGGGTGGGCTGGATGCGGAGTAAAGTTCTTTTGGCTCTTAAAGANAAGG 1140
```


CURRENT FILING DATE: 2020-11-28			
NUMBER OF SEQ ID NOS: 97222			
SOFTWARE: PatentIn version 3.2			
SEQ ID NO 34597			
LENGTH: 2391			
TYPE: DNA			
ORGANISM: Homo sapiens			
US-09-724-676-34597			
Query Match			
Best Local Similarity 98.0%; Pred. No. 2,8e+275;			
Matches 1139; Conservative 0; Mismatches 23; Indels 0; Gaps			
QY	1	GGGGGGGGGGGACCACTTGGCTGCTCCGTCGGCCGCGCACTTGGCTGCTCCGCT	60
Db	2391	GGGGGGGGGGGACCACTTGGCTGCTCCGTCGGCCGCGCGCACTTGGCTGCTCCGCT	2332
QY	61	CCCGCCGCGCACTTCCCTGCTCCGCTCCCGCCGCGCGCGCGCGCGCGCGCT	120
Db	2331	CCCGCCGCGCACTTCCCTGCTCCGCTCCCGCCGCGCGCGCGCGCGCGCGCT	2272
QY	121	TCGAGAGCTGCT	180
Db	2271	TCGAGAGCTGCT	2212
QY	181	CCGCT	240
Db	2211	CCGCT	2152
QY	241	AGGAGAGCGACGCGACGCGCGACCGCGACCGCTGCGGATTCGCGATCGAGCGCGCTAC	300
Db	2151	AGGAGAGCGACGCGCGACCGCGCGACCGCGACCGCTGCGGATTCGCGATCGAGCGCGCTAC	2092
QY	301	AGCTGAGAGATGAATTCCTCTGCTGACACCAACGTGTCTTGGAGGGGTGTTTG	360
Db	2091	AGCTGAGAGATGAATTCCTCTGCTGACACCAACGTGTCTTGGAGGGGTGTTTG	2032
QY	361	GAGAGTTGCTGTGTTTATCAAGGATTCACAAATGCTAAAGCTGTCTTCCAGGGA	420
Db	2031	GAGAGTTGCTGTGTTTATCAAGGATTCACAAATGCTAAAGCTGTCTTCCAGGGA	1972
QY	421	GAGAAATCTGGGATTCGCAATGATTCGCGAGCTTTTGGACGCGTGGATTCGACCC	480
Db	1971	GAGAAATCTGGGATTCGCAATGATTCGCGAGCTTTTGGACGCGTGGATTCGACCC	1912
QY	481	AGAGAGAGGGGAGCTTGGGCGCCAGTTTATGCTCTCAAGTGAGGCAATTTGGGGACGAA	540
Db	1911	AGAGAGAGGGGAGCTTGGGCGCCAGTTTATGCTCTCAAGTGAGGCAATTTGGGGACGAA	1852
QY	541	TACAGAGATATGGAATCAGATTTATTCAGACGAGGAGTTGACCACTCGCAAGAGTAAAT	600
Db	1851	TACAGAGATATGGAATCAGATTTATTCAGACGAGGAGTTGACCACTCGCAAGAGTAAAT	1792
QY	601	GACACCATCAAAAACAACCTCGACGACAGAGAAATCATATGTGGCTTGGAAATCCAGA	660
Db	1791	GACACCATCAAAAACAACCTCGACGACAGAGAAATCATATGTGGCTTGGAAATCCAGA	1732
QY	661	GATCTTCTCTGATGGCGTGTCTTCACATGCGCATGTCCTTCGCAAGTTCTATGTGGAA	720
Db	1731	GATCTTCTCTGATGGCGTGTCTTCACATGCGCATGTCCTTCGCAAGTTCTATGTGGAA	1672
QY	721	AGTGAAGCTGTCGCGCAGCTGACACAGATTCGGGAGCATGGGCGTGGGTGCTTTG	780
Db	1671	AGTGAAGCTGTCGCGCAGCTGACACAGATTCGGGAGCATGGGCGTGGGTGCTTTG	1612
QY	781	AACATGCGCGATGCGCGCTGCTGACGATCAATGATGCGAGCATCGCGCGCTGAACCA	840
Db	1611	AACATGCGCGATGCGCGCTGCTGACGATCAATGATGCGAGCATCGCGCGCTGAACCA	1552
QY	841	GGTGAATTATACACTTGGGAGATGACATATTTACCTGAATACATGAGACCACTG	900
Db	1551	GGTGAATTATACACTTGGGAGATGACATATTTACCTGAATACATGAGACCACTG	1492

Qy	1	GGGGGGGGGGGGGACACTTTGGCCGTGGCCCTGGCCGGGCGGCGACACTTGGGCTGGCCGT	60
Db	2391	GGGGGGGGGGGGGACACTTGGCTGGCTCTGGTCGGCGCGCGCACTTGGCTGGCTCGGT	2392
Qy	61	CCCGCCCGCGCACTTGGCTGGCTCGCTCCCGCCCGCCCGCGCATGCTGGGCGCGC	120
Db	2331	CCCGCCCGCGCACTTGGCTGGCTCGCTCCCGCCCGCCCGCGCATGCTGGGCGCGC	2272
Qy	121	TGGAGCTGCGCGCGCGCGCCTTGGCCCCCGCGCGCAGAGAGCGGAGCGCGAGCGCGT	180
Db	2271	TGGAGCTGCGCGCGCGCGCCTTGGCCCCCGCGCGCAGAGAGCGGAGCGCGAGCGGT	2212
Qy	181	CCGCGCGACGGGGAGCTGAGTACCTGGGGCAGATGCACAATCTCCGCTCGCGCGTC	240
Db	2211	CCGCGCGACGGGGAGCTGAGTACCTGGGGGAGATGCACAATCTCCGCTCGCGCGTC	2152
Qy	241	AGGAGAGCGACCGCACCGGGCACCGGCACCTCTGCGTATTGGGATGCGAGCGCGCTAC	300
Db	2151	AGGAGAGCGACCGCACCGGGCACCGGCACCTCTGCGTATTGGGATGCGAGCGCGCTAC	2092
Qy	301	AGCTCGAGAGTAAATTCCTCGCTGGTGAACAACCAACGATGTTCTGGAAAGGATGTTTG	360
Db	2091	AGCTCGAGAGTAAATTCCTCTGCTGCAACAACAACGATGTTCTGGAAAGGATGTTTG	2032
Qy	361	GAGGAGTGTGCTGTATTTCAGGAGATCCAAATGCTAAAGAGCTGTCTCCAAAGGCA	420
Db	2031	GAGGAGTGTGCTGTATTTCAGAGGATCCAAATGCTAAAGAGCTGTCTCCAAAGGCA	1972
Qy	421	GTTGAATTCGGAGTGCCTAATGGATCCCGAGACTTTTGGACAGCTGGGATTCCTACAC	480
Db	1971	GTTGAATTCGGAGTGCCTAATGGATCCCGAGACTTTTGGACAGCTGGGATTCCTACAC	1912

Query Match	73.3%;	Score 1126.4;	DB 5;	Length 2411;
Best Local Similarity	98.0%;	Pred. No. 2.8e-275;		

Matches 1139; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

```
OY 1 GGGGGGGGGGGGACCACTTGGCTGCTCCGTCGCGCCGCACTTGGCTGCTCCGTC
Db 2411 GGGGGGGGGGGGACCACTTGGCTGCTCCGTCGCGCCGCACTTGGCTGCTCCGTC
OY 61 CCGCGCGCGCACTTGGCTGCTCCGTCGCGCCGCACTTGGCTGCTCCGTC
Db 2351 CCGCGCGCGCACTTGGCTGCTCCGTCGCGCCGCACTTGGCTGCTCCGTC
OY 121 TCGGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
Db 2291 TCGGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
OY 181 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
Db 2231 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
OY 241 AGGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
Db 2171 AGGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
OY 301 AGCGTGAAGATGAATTCCTCTGCTGCAACCAACGCTGCTGCAAGGGGCTTTG
Db 2111 AGCGTGAAGATGAATTCCTCTGCTGCAACCAACGCTGCTGCAAGGGGCTTTG
OY 361 GAGGAGTGTGCTGCTGCTTATCAAGGATCCCAAAATGCTTAAAGAGCTGCTTCAAGGGA
Db 2051 GAGGAGTGTGCTGCTGCTTATCAAGGATCCCAAAATGCTTAAAGAGCTGCTTCAAGGGA
OY 421 GTGAAATCTGGGATGCCAATGGAATCCCGAGACTTTTGAACGCTGCGGATTTCCACC
Db 1991 GTGAAATCTGGGATGCCAATGGAATCCCGAGACTTTTGAACGCTGCGGATTTCCACC
OY 481 AGGAGGAGGGGAGCTTGGGCGCGAGTTTATGCTTCAAGGAGCTTGGGCGAGAA
Db 1931 AGGAGGAGGGGAGCTTGGGCGCGAGTTTATGCTTCAAGGAGCTTGGGCGAGAA
OY 541 TACAGAGATATGGAATCAGATTATTCAGGACAGGAGGTGATCCAACTGCAAAAGTAT
Db 1871 TACAGAGATATGGAATCAGATTATTCAGGACAGGAGGTGATCCAACTGCAAAAGTAT
OY 601 GACACATCAAAAACCAACCTTGAAGCAAGAAATCATCATGTGCGCTTGAATTCAGAA
Db 1811 GACACATCAAAAACCAACCTTGAAGCAAGAAATCATCATGTGCGCTTGAATTCAGAA
OY 661 GATTTCTCTGATGAGCGCTGCTCCATGCGCATGCCCTCTGCGAGTTGATGCTGTAAC
Db 1751 GATTTCTCTGATGAGCGCTGCTCCATGCGCATGCCCTCTGCGAGTTGATGCTGTAAC
OY 721 AGTGAAGCTGCTGCGAGCTGATCAAGAGATCGGAGAGCATGGGCGCTGGTGTCTTTC
Db 1691 AGTGAAGCTGCTGCGAGCTGATCAAGAGATCGGAGAGCATGGGCGCTGGTGTCTTTC
OY 781 AACATCGCGAGCTAGCGCTGCTCACTGATCATGTTGCGCAATCAACGCGCGCTGAAGCCA
Db 1631 AACATCGCGAGCTAGCGCTGCTCACTGATCATGTTGCGCAATCAACGCGCGCTGAAGCCA
OY 841 GGTGACTTTATACACTTTGGGAGATGCAATATTTACCTGATTCACATGAGCCACTG
Db 1571 GGTGACTTTATACACTTTGGGAGATGCAATATTTACCTGATTCACATGAGCCACTG
OY 901 AAAATTCAGCTTACGAGGAGAACCAAGACTTTTCCAAAGCTGAGATTTCTTCAAAAGTT
Db 1511 AAAATTCAGCTTACGAGGAGAACCAAGACTTTTCCAAAGCTGAGATTTCTTCAAAAGTT
OY 961 GAGGAAATGATGACTTCAAAAGCTGAGAACCTTTCAGATTGAGGGGTACAATCCGATCCA
Db 1451 GAGGAAATGATGACTTCAAAAGCTGAGAACCTTTCAGATTGAGGGGTACAATCCGATCCA
OY 1021 ACTATTTAAATGGAATGCTGCTTATGAGGCTCTTCAAAAGGAGCTGGAAGGATATGCA
Db 1391 ACTATTTAAATGGAATGCTGCTTATGAGGCTCTTCAAAAGGAGCTGGAAGGATATGCA
```

```
OY 1081 GTCTTTAGGGGTTGGGCTGATGCCAGTAAAGTTCTTTTGTCTTAAAGANAAG
Db 1331 GTCTTTAGGGGTTGGGCTGATGCCAGTAAAGTTCTTTTGTCTTAAAGANAAG
OY 1141 AACTAGTCAAAAATCTGCTCG 1162
Db 1271 AGGGGTATGACTTCTAGCTG 1250
```

RESULT 10

```
US-09-724-676A-34535/c
; Sequence 34535, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 34535
; LENGTH: 2411
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURES:
; NAME/KEY: misc feature
; LOCATION: (41)-(41)
; OTHER INFORMATION: n is a,c,g, or t
US-09-724-676A-34535
```

Query Match 73.3%; Score 1126.4; DB 5; Length 2411;

Best Local Similarity 98.0%; Pred. No. 2,8e-275;

Matches 1139; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

```
OY 1 GGGGGGGGGGGGACCACTTGGCTGCTCCGTCGCGCCGCACTTGGCTGCTCCGTC
Db 2411 GGGGGGGGGGGGACCACTTGGCTGCTCCGTCGCGCCGCACTTGGCTGCTCCGTC
OY 61 CCGCGCGCGCACTTGGCTGCTCCGTCGCGCCGCACTTGGCTGCTCCGTC
Db 2351 CCGCGCGCGCACTTGGCTGCTCCGTCGCGCCGCACTTGGCTGCTCCGTC
OY 121 TCGGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
Db 2291 TCGGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
OY 181 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
Db 2231 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
OY 241 AGGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
Db 2171 AGGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
OY 301 AGCGTGAAGATGAATTCCTCTGCTGCAACCAACGCTGCTGCAAGGGGCTTTG
Db 2111 AGCGTGAAGATGAATTCCTCTGCTGCAACCAACGCTGCTGCAAGGGGCTTTG
OY 361 GAGGAGTGTGCTGCTGCTTATCAAGGATCCCAAAATGCTTAAAGAGCTGCTTCAAGGGA
Db 2051 GAGGAGTGTGCTGCTGCTTATCAAGGATCCCAAAATGCTTAAAGAGCTGCTTCAAGGGA
OY 421 GTGAAATCTGGGATGCCAATGGAATCCCGAGACTTTTGAACGCTGCGGATTTCCACC
Db 1991 GTGAAATCTGGGATGCCAATGGAATCCCGAGACTTTTGAACGCTGCGGATTTCCACC
OY 481 AGGAGGAGGGGAGCTTGGGCGCGAGTTTATGCTTCAAGGAGCTTGGGCGAGAA
Db 1931 AGGAGGAGGGGAGCTTGGGCGCGAGTTTATGCTTCAAGGAGCTTGGGCGAGAA
OY 541 TACAGAGATATGGAATCAGATTATTCAGGACAGGAGGTGATCCAACTGCAAAAGTAT
```

Db 1871 TACAGAGATATGATAGATTATTCAGGACAGGAGCTGACCAACTGCAAAAGTGATT 1812
Qy 601 GACACATCAAAACCAACCTTGACAGACAGAAATATCATCTGTGAGCTTTGAAATCAGA 650
Db 1811 GACACATCAAAACCAACCTTGACAGACAGAAATATCATCTGTGAGCTTTGAAATCAGA 1752
Qy 661 GATCTTCTCTGATGAGGCTGCTCCATCCATCCCTCTGACAGTTCTATGTGTGAC 720
Db 1751 GATCTTCTCTGATGAGGCTGCTCCATCCATCCCTCTGACAGTTCTATGTGTGAC 1692
Qy 721 AGTGAAGTCTCTGACCTGCTGACAGAGATCGGAGACATGAGCTCGGTGCTCTTC 780
Db 1691 AGTGAAGTCTCTGACCTGCTGACAGAGATCGGAGACATGAGCTCGGTGCTCTTC 1632
Qy 781 AACATCGCAGCTTACGCTCTGCTCACTGATGATGTCGACATCAAGGCTTAAACCA 840
Db 1631 AACATCGCAGCTTACGCTCTGCTCACTGATGATGTCGACATCAAGGCTTAAACCA 1572
Qy 841 GGTGACTTTATACACTTTGGGAGATGACATATTACCTGATCATCGAGCACTG 900
Db 1571 GGTGACTTTATACACTTTGGGAGATGACATATTACCTGATCATCGAGCACTG 1512
Qy 901 AAAATTCAGCTTCAGCAGACCAAGACCTTCCAAAGCTCAGATTCCTTGAAGATT 960
Db 1511 AAAATTCAGCTTCAGCAGACCAAGACCTTCCAAAGCTCAGATTCCTTGAAGATT 1452
Qy 961 GAGAAATTTGATGACTTCAAGCTGAAGACTTTCATGATGAGGTTAATTCGACATCA 1020
Db 1451 GAGAAATTTGATGACTTCAAGCTGAAGACTTTCATGATGAGGTTAATTCGACATCA 1392
Qy 1021 ACTATTAAAAATGAAATGGCTGTTTAAAGGCTCTTCAAGGAGCTTGAAGATTTTCA 1080
Db 1391 ACTATTAAAAATGAAATGGCTGTTTAAAGGCTCTTCAAGGAGCTTGAAGATTTTCA 1332
Qy 1081 GTCTTAAAGGCTTGGGCTGATGACGAGGTTAAAGTTCTTTTGTCTTAAAGAAAG 1140
Db 1331 GTCTTAAAGGCTTGGGCTGATGACGAGGTTAAAGTTCTTTTGTCTTAAAGTTCTG 1272
Qy 1141 AACTAGTCAAAAATCTGTCCG 1162
Db 1271 AGGCTGATGACTTCTAGCTG 1250

RESULT 11
US-09-724-676-34589/c
; Sequence 34589, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 34589
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676-34589

Query Match 73.3% Score 1126.4; DB 5; Length 2491;
Best Local Similarity 98.0%; Pred. No. 2,9e-275;
Matches 1139; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGACCACTTGGCTTCCGCTCCGCGCGGCACTTGGCTTCCGCT 60
Db 2491 GGGGGGGGGGGGACCACTTGGCTTCCGCTCCGCGCGGCACTTGGCTTCCGCT 2432
Qy 61 CCGCGCGGCACTTGGCTTCCGCGCGGCTCCGCGCGGCACTTGGCTTCCGCT 120
Db 2431 CCGCGCGGCACTTGGCTTCCGCGCGGCTCCGCGCGGCACTTGGCTTCCGCT 2372
Qy 121 TCGAGAGCTGCGCGCGCTTGCCTCCCGCGGCAAGAGCGGAGCGCGAGCTGCT 180

Db 2371 TCGAGAGCTGCGCGCGCTTGCCTCCCGCGGCAAGAGCGGAGCGCGAGCTGCT 2212
Qy 181 CCGCGCGGCACTTGGCTTCCGCGCGGCTCCGCGCGGCACTTGGCTTCCGCT 240
Db 2311 CCGCGCGGCACTTGGCTTCCGCGCGGCTCCGCGCGGCACTTGGCTTCCGCT 2252
Qy 241 AGAAGAGAGCACTGACCGGAGCAAGGACCTTGGCTTGGAGATGAGGAGCTGAC 300
Db 2251 AGAAGAGAGCACTGACCGGAGCAAGGACCTTGGCTTGGAGATGAGGAGCTGAC 2192
Qy 301 AGCTGAGAGATGATTCCTCTGCTGACCAACCAAGTGTCTTGGAGAGGCTTTG 360
Db 2191 AGCTGAGAGATGATTCCTCTGCTGACCAACCAAGTGTCTTGGAGAGGCTTTG 2132
Qy 361 GAGGAGTGTCTGTGTTATCAAGGATCCAATGCTTAAAGCTTCTTCAAGGAG 420
Db 2131 GAGGAGTGTCTGTGTTATCAAGGATCCAATGCTTAAAGCTTCTTCAAGGAG 2072
Qy 421 GTGAAAATCTGGGATGCCATGATATCCGAGACTTTTGGACAGCTTGGAGTTCTCAC 480
Db 2071 GTGAAAATCTGGGATGCCATGATATCCGAGACTTTTGGACAGCTTGGAGTTCTCAC 2012
Qy 481 AGAAGAGAGGAGCTTGGGACCAAGTTATGCTTCAGTGGAGCACTTTGGGAGAG 540
Db 2011 AGAAGAGAGGAGCTTGGGACCAAGTTATGCTTCAGTGGAGCACTTTGGGAGAG 1952
Qy 541 TACAGATATGAAATCAAGTTATTCAGGACAGGAGTGAACAATGCAAGAGTAT 600
Db 1951 TACAGATATGAAATCAAGTTATTCAGGACAGGAGTGAACAATGCAAGAGTAT 1892
Qy 601 GACACATCAAAACCAACCTTGACAGACAGAAATATCATCTGTGAGCTTGAATCCA 660
Db 1891 GACACATCAAAACCAACCTTGACAGACAGAAATATCATCTGTGAGCTTGAATCCA 1832
Qy 661 GATCTTCTCTGATGAGGCTGCTCCATCCATCCCTCTGACAGTTCTATGTGTGAC 720
Db 1831 GATCTTCTCTGATGAGGCTGCTCCATCCATCCCTCTGACAGTTCTATGTGTGAC 1772
Qy 721 AGTGAAGTCTCTGACCTGCTGACAGAGATCGGAGACATGAGGCTTGAAGATT 780
Db 1771 AGTGAAGTCTCTGACCTGCTGACAGAGATCGGAGACATGAGGCTTGAAGATT 1712
Qy 781 AACATCGCAGCTTACGCTCTGCTCACTGATGATGTCGACATCAAGGCTTAAACCA 840
Db 1711 AACATCGCAGCTTACGCTCTGCTCACTGATGATGTCGACATCAAGGCTTAAACCA 1652
Qy 841 GGTGACTTTATACACTTTGGGAGATGACATATTACCTGATCATCGAGCACTG 900
Db 1651 GGTGACTTTATACACTTTGGGAGATGACATATTACCTGATCATCGAGCACTG 1592
Qy 901 AAAATTCAGCTTCAGCAGACCAAGACCTTCCAAAGCTCAGATTCCTTGAAGATT 960
Db 1591 AAAATTCAGCTTCAGCAGACCAAGACCTTCCAAAGCTCAGATTCCTTGAAGATT 1532
Qy 961 GAGAAATTTGATGACTTCAAGCTGAAGACTTTCATGATGAGGTTAATTCGACATCA 1020
Db 1531 GAGAAATTTGATGACTTCAAGCTGAAGACTTTCATGATGAGGTTAATTCGACATCA 1472
Qy 1021 ACTATTAAAAATGAAATGGCTGTTTAAAGGCTCTTCAAGGAGCTTGAAGATTTTCA 1080
Db 1471 ACTATTAAAAATGAAATGGCTGTTTAAAGGCTCTTCAAGGAGCTTGAAGATTTTCA 1412
Qy 1081 GTCTTAAAGGCTTGGGCTGATGACGAGGTTAAAGTTCTTTTGTCTTAAAGAAAG 1140
Db 1411 GTCTTAAAGGCTTGGGCTGATGACGAGGTTAAAGTTCTTTTGTCTTAAAGTTCTG 1352
Qy 1141 AACTAGTCAAAAATCTGTCCG 1162
Db 1351 AGGCTGATGACTTCTAGCTG 1330

RESULT 12

```

US-09-724-676A-34589/c
: Sequence 34589, Application US/09724676A
: GENERAL INFORMATION:
: APPLICANT: Compugen LTD
: TITLE OF INVENTION: Variants of alternative splicing
: FILE REFERENCE: 129181.4 Compugen
: CURRENT APPLICATION NUMBER: US/09/724,676A
: NUMBER OF SEQ ID NOS: 97222
: SOFTWARE: PatentIn version 3.2
: SEQ ID NO 34589
: LENGTH: 2491
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-724-676A-34589

Query Match          73.3%; Score 1126.4; DB 5; Length 2491;
Best Local Similarity 98.0%; Pred. No. 2.9e-275;
Matches 1139; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 1 GGGGGGGGGGGGACCACTTGGCTGCTCCGTCGCGCGCACTTGGCTGCTCCGCT 60
Db 2491 GGGGGGGGGGGGACCACTTGGCTGCTCCGTCGCGCGCACTTGGCTGCTCCGCT 2432
QY 61 CCGCGCGCGCACTTGGCTGCTCCGTCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
Db 2431 CCGCGCGCGCACTTGGCTGCTCCGTCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2372
QY 121 TCGGAGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
Db 2371 TCGGAGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2312
QY 181 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240
Db 2311 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2252
QY 241 AGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
Db 2251 AGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2192
QY 301 AGCCTGAGAGATGATTCCTCTGCTGACCAACCAACGATGCTTGGAAAGGATG 360
Db 2191 AGCCTGAGAGATGATTCCTCTGCTGACCAACCAACGATGCTTGGAAAGGATG 2132
QY 361 GAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
Db 2131 GAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2072
QY 421 GTGAAATATGAGATGAGATGATGATGATGATGATGATGATGATGATGATGAT 480
Db 2071 GTGAAATATGAGATGAGATGATGATGATGATGATGATGATGATGATGATGAT 2012
QY 481 AGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
Db 2011 AGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1952
QY 541 TACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
Db 1951 TACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1892
QY 601 GACACCATCAAAACCAACCTTGCAGACAGAGAGATGATGATGATGATGATGAT 660
Db 1891 GACACCATCAAAACCAACCTTGCAGACAGAGAGATGATGATGATGATGATGAT 1832
QY 661 GATCTTCTCTGATGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
Db 1831 GATCTTCTCTGATGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1772
QY 721 AGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db 1771 AGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1712
QY 781 AACATGCGACGTAAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840

```

```

Db 1711 AACATGCGACGTAAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1652
QY 841 GGTGAATTAATACACACTTGGGAGATGACATATTATCTGATCAATCGACCACTG 900
Db 1651 GGTGAATTAATACACACTTGGGAGATGACATATTATCTGATCAATCGACCACTG 1592
QY 901 AAAATTCACTTCAAGCGAGAACCAACCTTCCCAAGCTCAGGATTTCTGAAAAGTT 960
Db 1591 AAAATTCACTTCAAGCGAGAACCAACCTTCCCAAGCTCAGGATTTCTGAAAAGTT 1532
QY 961 GAGAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
Db 1531 GAGAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1472
QY 1021 ACTATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
Db 1471 ACTATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1412
QY 1081 GTCTTTAAGGGGTTGGGCTGAGATGCGAGGTTAAAGTTCTTTTGTCTTAAAGAAAAG 1140
Db 1411 GTCTTTAAGGGGTTGGGCTGAGATGCGAGGTTAAAGTTCTTTTGTCTTAAAGAAAAG 1352
QY 1141 AACTAGATCAAAAATCTGTCCG 1162
Db 1351 AAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1330

RESULT 13
US-09-724-676-34548/c
: Sequence 34548, Application US/09724676
: GENERAL INFORMATION:
: APPLICANT: Compugen LTD
: TITLE OF INVENTION: Variants of alternative splicing
: FILE REFERENCE: 129181.4 Compugen
: CURRENT APPLICATION NUMBER: US/09/724,676
: NUMBER OF SEQ ID NOS: 97222
: SOFTWARE: PatentIn version 3.2
: SEQ ID NO 34548
: LENGTH: 2527
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc.feature
: LOCATION: (41)..(41)
: OTHER INFORMATION: n is a,c,g, or t
US-09-724-676-34548

Query Match          73.3%; Score 1126.4; DB 5; Length 2527;
Best Local Similarity 98.0%; Pred. No. 2.9e-275;
Matches 1139; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 1 GGGGGGGGGGGGACCACTTGGCTGCTCCGTCGCGCGCACTTGGCTGCTCCGCT 60
Db 2527 GGGGGGGGGGGGACCACTTGGCTGCTCCGTCGCGCGCGCGCGCGCGCGCGCGCG 2468
QY 61 CCGCGCGCGCACTTGGCTGCTCCGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
Db 2467 CCGCGCGCGCACTTGGCTGCTCCGTCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2408
QY 121 TCGGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
Db 2407 TCGGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2348
QY 181 CCGCGCGCGAGAGCTGAGTACCTGAGGAGATGATCAACATCTCTCGCTGCGCGCTC 240
Db 2347 CCGCGCGCGAGAGCTGAGTACCTGAGGAGATGATCAACATCTCTCGCTGCGCGCTC 2288
QY 241 AGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
Db 2287 AGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2228

```

```

QY 301 AGCTGAGATGAATTCCTTCTGCTGACACCAACGCTGTGTTTGGAGGCTGTTTGG 360
D 2227 AGCTGAGAGGTGAATTCCTTCTGCTGACACCAACGCTGTGTTTGGAGGCTGTTTGG 2168
QY 361 GAGAGTGTGCTGTGTTTATCAAGGATCCCAAAATGTTAAAGAGCTGTCTTCAAGGA 420
D 2167 GAGAGTGTGCTGTGTTTATCAAGGATCCCAAAATGTTAAAGAGCTGTCTTCAAGGA 2108
QY 421 GTGAAAATCTGGGATGCCAATGATGATCCGAGACTTTTGGACAGCTGGGATCTCCAC 480
D 2107 GTGAAAATCTGGGATGCCAATGATGATCCGAGACTTTTGGACAGCTGGGATCTCCAC 2048
QY 481 AGAGAGAGGAGCTTGGGCGCCAGTTTATGCTTCAAGTGAAGCTTTTGGGCGAGA 540
D 2047 AGAGAGAGGAGCTTGGGCGCCAGTTTATGCTTCAAGTGAAGCTTTTGGGCGAGA 1988
QY 541 TACAGAGATATGAAATCAGATTATTCAGAGACAGGAGTTGACCACTGCAAGAGTATT 600
D 1987 TACAGAGATATGAAATCAGATTATTCAGAGACAGGAGTTGACCACTGCAAGAGTATT 1928
QY 601 GACACATCAAAACCAACCTGACGACAGAGATATCATATGTCGCTTGGAAATCCAGA 660
D 1927 GACACATCAAAACCAACCTGACGACAGAGATATCATATGTCGCTTGGAAATCCAGA 1868
QY 661 GATCTTCTCTGATGAGGCTGCTCCATGCAAGCCCTGCGACATATGATGATGATGAC 720
D 1867 GATCTTCTCTGATGAGGCTGCTCCATGCAAGCCCTGCGACATATGATGATGATGAC 1808
QY 721 AGTACGCTGCTCTGACAGCTGTACAGAGATCGGAGACATGAGGCTCGTGTGCTTTC 780
D 1807 AGTACGCTGCTCTGACAGCTGTACAGAGATCGGAGACATGAGGCTCGTGTGCTTTC 1748
QY 781 AACATGCGACGCTACGCTCTGCTACAGTACATGATGCGACATACGAGGCTGAAAGCA 840
D 1747 AACATGCGACGCTACGCTCTGCTACAGTACATGATGCGACATACGAGGCTGAAAGCA 1688
QY 841 GGTGACTTTATACACTTTGGAGATGACATATTTTCCGAAATCATACAGGCACTG 900
D 1687 GGTGACTTTATACACTTTGGAGATGACATATTTTCCGAAATCATACAGGCACTG 1628
QY 901 AAAATTCAGCTTACGAGAACCCAGACCTTCCGAAAGCTCAGAGATCTTGAAAGTT 960
D 1627 AAAATTCAGCTTACGAGAACCCAGACCTTCCGAAAGCTCAGAGATCTTGAAAGTT 1568
QY 961 GAGAAATGATGATCTTCAAGCTGAAGCTTTCAGATTTAGAGGATCAATCCGATCCA 1020
D 1567 GAGAAATGATGATCTTCAAGCTGAAGCTTTCAGATTTAGAGGATCAATCCGATCCA 1508
QY 1021 ACTATTAATAATGAAATGCTGTTTAAAGGCTCTTCAAGAGCTTGAAGATATTGCA 1080
D 1507 ACTATTAATAATGAAATGCTGTTTAAAGGCTCTTCAAGAGCTTGAAGATATTGCA 1448
QY 1081 GTCTTTAAGGAGTGTGGCTGATATGCGAGGATAAGTCTTTTGTCTTAAGAAANAAG 1140
D 1447 GTCTTTAAGGAGTGTGGCTGATATGCGAGGATAAGTCTTTTGTCTTAAGAAANAAG 1388
QY 1141 AACTAGTCAAAAATCTGTCCG 1162
D 1387 AAGGTGTATGACTTCTAGCTG 1366

```

RESULT 14

```

US-09-724-676A-34548/c
; Sequence 34548, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 34548

```

```

; LENGTH: 2527
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (41)..(41)
; OTHER INFORMATION: n is a,c,g, or t
US-09-724-676A-34548

Query Match 73.3%; Score 1126.4; DB 5; Length 2527;
Best Local Similarity 98.0%; Pred. No. 2,9e-275;
Matches 1139; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 1 GGGGGGGGGGGGGGACCACTTGGCTGCTCCGCTCCGCGCGGCACTTGGCTGCTCCGCT 60
D 2527 GGGGGGGGGGGGGGACCACTTGGCTGCTCCGCTCCGCGCGGCACTTGGCTGCTCCGCT 2468
QY 61 CCGCGCGCGCGCACTTGGCTGCTCCGCTCCGCGCGGCACTTGGCTGCTCCGCT 120
D 2467 CCGCGCGCGCGCACTTGGCTGCTCCGCTCCGCGCGGCACTTGGCTGCTCCGCT 2408
QY 121 TGGAGCTGCTGCGCGCGCGCTTCCGCGCGGCACTTGGCTGCTCCGCT 180
D 2407 TGGAGCTGCTGCGCGCGCGCTTCCGCGCGGCACTTGGCTGCTCCGCT 2348
QY 181 CCGCGCGCGCGGAGCTGCACTTGGGCGAGATCCACACATCTCTCGCTGCGCGCT 240
D 2347 CCGCGCGCGCGGAGCTGCACTTGGGCGAGATCCACACATCTCTCGCTGCGCGCT 2288
QY 241 AGAGAGAGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
D 2287 AGAGAGAGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2228
QY 301 AGCTGAGATGAATTCCTGCTGCTGCAACAACGATGTTTGGAGAGGATGTTTGG 360
D 2227 AGCTGAGATGAATTCCTGCTGCTGCAACAACGATGTTTGGAGAGGATGTTTGG 2168
QY 361 GAGAGTGTGCTGTGTTTATCAAGGATCCCAAAATGCTTAAAGCTGTCTTCAAGGGA 420
D 2167 GAGAGTGTGCTGTGTTTATCAAGGATCCCAAAATGCTTAAAGCTGTCTTCAAGGGA 2108
QY 421 GTGAAAATCTGGGATGCCAATGATGATCCGAGACTTTTGGACAGCTGGGATCTCCAC 480
D 1987 GTGAAAATCTGGGATGCCAATGATGATCCGAGACTTTTGGACAGCTGGGATCTCCAC 1928
QY 481 GACACATCAAAACCAACCTGACGACAGAGATATCATATGTCGCTTGGAAATCCAGA 660
D 1927 GACACATCAAAACCAACCTGACGACAGAGATATCATATGTCGCTTGGAAATCCAGA 1868
QY 661 GATCTTCTCTGATGAGGCTGCTCCATGCAAGCCCTGCGACATATGATGATGATGAC 720
D 1867 GATCTTCTCTGATGAGGCTGCTCCATGCAAGCCCTGCGACATATGATGATGATGAC 1808
QY 721 AGTACGCTGCTCTGACAGCTGTACAGAGATCGGAGACATGAGGCTCGTGTGCTTTC 780
D 1807 AGTACGCTGCTCTGACAGCTGTACAGAGATCGGAGACATGAGGCTCGTGTGCTTTC 1748
QY 781 AACATGCGACGCTACGCTCTGCTACAGTACATGATGCGACATACGAGGCTGAAAGCA 840
D 1747 AACATGCGACGCTACGCTCTGCTACAGTACATGATGCGACATACGAGGCTGAAAGCA 1688
QY 841 GGTGACTTTATACACTTTGGAGATGACATATTTTCCGAAATCATACAGGCACTG 900
D 1687 GGTGACTTTATACACTTTGGAGATGACATATTTTCCGAAATCATACAGGCACTG 1628

```

```

QY 901 AAAATTAGCTTACGAGAACCCGACCTTTCCCAAAGCTTCGAGATCTTCCAAAGTT 960
DB 1627 AAAATTCAGCTTCAGGAGAACCCGACCTTTCCCAAAGCTTCGAGATCTTCCAAAGTT 1568
QY 961 GAGAAATTTAGTACCTTCAAGCTGAAGATCTTCAGATTTGAAGGGGTAACATCCGATCCA 1020
DB 1567 GAGAAATTTAGTACCTTCAAGCTGAAGATCTTCAGATTTGAAGGGGTAACATCCGATCCA 1508
QY 1021 ACTATTAATGGAATGAGCTGTTTAGGGTCTTTCAAAGAGCTTGAAGGATATTTGCA 1080
DB 1507 ACTATTAATGGAATGAGCTGTTTAGGGTCTTTCAAAGAGCTTGAAGGATATTTGCA 1448
QY 1081 GTCTTTAGGGGTTGGCTGGATGCCAGAGTAAAGTTCTTTTCTCTPAAAGAAAG 1140
DB 1447 GTCTTTAGGGGTTGGCTGGATGCCAGAGTAAAGTTCTTTTCTCTTAAAGATTCCTG 1388
QY 1141 AACTAGGTCAAAAATCTGTCCG 1162
DB 1387 AGGGGTATGACTTCTAGCTG 1366

RESULT 15
US-09-724-676-34565/c
; Sequence 34565, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 34565
; LENGTH: 3023
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676-34565

Query Match          73.3%; Score 1126.4; DB 5; Length 3023;
Best Local Similarity 98.0%; Pred. No. 3.1e-275;
Matches 1139; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 1 GGGGGGGGGGGGACCACTTGCGCTTCCCTCCCGCGGCACTTGCGCTTCCCT 60
DB 3023 GGGGGGGGGGGGACCACTTGCGCTTCCCTCCCGCGGCACTTGCGCTTCCCT 2964
QY 61 CCCGCCCGCCACTTGGCTGCTGCTCCGCTCCCGCGGCACTTGCGCTTCCCT 120
DB 2963 CCCGCCCGCCACTTGGCTGCTGCTCCGCTCCCGCGGCACTTGCGCTTCCCT 2904
QY 121 TCGAGGTGCGCGCGCGGCGCTTGGCCCGCGGCACTTGCGCGGCGGCGT 180
DB 2903 TCGAGGTGCGCGCGCGGCGCTTGGCCCGCGGCACTTGCGCGGCGGCGT 2844
QY 181 CCGCGCGACGGGAGCTGACATCTGAGGAGTCCAAACATCTTCCGTCGCGCTC 240
DB 2843 CCGCGCGACGGGAGCTGACATCTGAGGAGTCCAAACATCTTCCGTCGCGCTC 2784
QY 241 AGAAGAGACGACCGACCGGCACTTGCGCTTCCGATTCGCAATGCAAGGCGGCTAC 300
DB 2783 AGAAGAGACGACCGACCGGCACTTGCGCTTCCGATTCGCAATGCAAGGCGGCTAC 2724
QY 301 AGCTTGAGAGATGAATTCCTCTGCTGACAAACATGCTGCTTGAAGGGTCTTTTG 360
DB 2723 AGCTTGAGAGATGAATTCCTCTGCTGACAAACATGCTGCTTGAAGGGTCTTTTG 2664
QY 361 GAGGAGTGTGCTGTGTTTCAAGGAGTCCAAATGCTTGAAGGAGTCTTCCAAAGGA 420
DB 2663 GAGGAGTGTGCTGTGTTTCAAGGAGTCCAAATGCTTGAAGGAGTCTTCCAAAGGA 2604
QY 421 GTGAAATCTGAGATGCCAATGATCCGAGACTTTTGAAGAGCTGAGATTCACAC 480
DB 2603 GTGAAATCTGAGATGCCAATGATCCGAGACTTTTGAAGAGCTGAGATTCACAC 2544

```

```

QY 481 AGAAGAAAGGGGACTTGGGCCGAGTTTATGCTTCAGATGAGGCAATTTTGGGCGAGA 540
DB 2543 AGAAGAAAGGGGACTTGGGCCGAGTTTATGCTTCAGATGAGGCAATTTTGGGCGAGA 2484
QY 541 TACAGAGATGGAATCAAGTATTCAGAGCAGGAGTTCACCACTGCAAAAGTGAAT 600
DB 2483 TACAGAGATGGAATCAAGTATTCAGAGCAGGAGTTCACCACTGCAAAAGTGAAT 2424
QY 601 GACACATCAAAACCAACCTGACGACGAGAGATATATATGCTGCTTGAATCCAAAG 660
DB 2423 GACACATCAAAACCAACCTGACGACGAGAGATATATATGCTGCTTGAATCCAAAG 2364
QY 661 GATTTCTCTGATGAGGCTGCTCCATGCGCATGCCCTGCACTTCTATGAGTGAAC 720
DB 2363 GATTTCTCTGATGAGGCTGCTCCATGCGCATGCCCTGCACTTCTATGAGTGAAC 2304
QY 721 AGTGAAGTGTCTGCGCAGCTGACGAGATCCGAGAGATGAGGAGATGAGGCTGCTTTC 780
DB 2303 AGTGAAGTGTCTGCGCAGCTGACGAGATCCGAGAGATGAGGAGATGAGGCTGCTTTC 2244
QY 781 AACATGCCAGTAAAGCCCTGCTCAGTACATGATTCGACATCAAGGCTGAGACCA 840
DB 2243 AACATGCCAGTAAAGCCCTGCTCAGTACATGATTCGACATCAAGGCTGAGACCA 2184
QY 841 GGTGACTTTATACACACTTGGAGATGACATATTTACTGATACATGAGGCTG 900
DB 2183 GGTGACTTTATACACACTTGGAGATGACATATTTACTGATACATGAGGCTG 2124
QY 901 AAAATTCAGCTTACGAGAACCCGACCTTTCCCAAAGCTCAGGATTTCTGAAAGTT 960
DB 2123 AAAATTCAGCTTACGAGAACCCGACCTTTCCCAAAGCTCAGGATTTCTGAAAGTT 2064
QY 961 GAGAAATTTAGTACCTTCAAGCTGAAGATCTTCAGATTTGAAGGGTAAATCCGATCCA 1020
DB 2063 GAGAAATTTAGTACCTTCAAGCTGAAGATCTTCAGATTTGAAGGGTAAATCCGATCCA 2004
QY 1021 ACTATTAATGGAATGAGCTGTTTAGGGTCTTTCAAAGAGCTTGAAGGATATTTGCA 1080
DB 2003 ACTATTAATGGAATGAGCTGTTTAGGGTCTTTCAAAGAGCTTGAAGGATATTTGCA 1944
QY 1081 GTCTTTAGGGGTTGGCTGGATGCCAGAGTAAAGTTCTTTTCTCTPAAAGAAAG 1140
DB 1943 GTCTTTAGGGGTTGGCTGGATGCCAGAGTAAAGTTCTTTTCTCTTAAAGATTCCTG 1884
QY 1141 AACTAGGTCAAAAATCTGTCCG 1162
DB 1883 AACTAGGTCAAAAATCTGTCCG 1862

```

Search completed: November 29, 2002, 00:20:06
Job time : 123.086 secs

Qy	1381	AAGCTATTTTGGAAATAATTTTTAGAATAATTTTAAGAAATTTCACAGCTATTCCTCCAAAT	1440
Db	1381	AAGCTATTTTGGAAATAATTTTAGAATAATTTTAAGAAATTTCACAGCTATTCCTCCAAAT	1440
Qy	1441	CTGAGGAGCTCAGTAACACCATCGATCATCATGTAGAGTGTTGAACCTTANAGT	1500
Db	1441	CTGAGGAGCTCAGTAACACCATCGATCATCATGTAGAGTGTTGAACCTTATAGT	1500
Qy	1501	TGTTTTATATGTTGCTATAATAAAGAGTGTTCTGC	1536
Db	1501	TGTTTTATATGTTGCTATAATAAAGAGTGTTCTGC	1536

RESULT 2
PCT-US99-12080-1
; Sequence 1, Application PC/TUS9912080
; GENERAL INFORMATION:
; APPLICANT: Dean, Nicholas M.
; TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDE MODULATION OF
; TITLE OF INVENTION: HUMAN THYMIDYLATE SYNTHASE EXPRESSION
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Jane Massey Licata
; STREET: 66 East Main Street
; CITY: Marlton
; STATE: NJ
; COUNTRY: USA
; ZIP: 08053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: WINDOWS 95
; SOFTWARE: WORD PERFECT 6.1 FOR WINDOWS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US99/12080
; FILING DATE: herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/089,195
; FILING DATE: June 2, 1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Jane Massey Licata
; REGISTRATION NUMBER: 32,257
; REFERENCE/DOCKET NUMBER: ISPH-0369
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (609) 810-1515
; TELEFAX: (609) 810-1454
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1536 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Unknown
; PUBLICATION INFORMATION:
; AUTHORS: Takeishi,K.
; AUTHORS: Kaneda,S.
; AUTHORS: Ayusawa,D.
; AUTHORS: Shimizu,K.
; AUTHORS: Gotoh,O.
; AUTHORS: Seno,T.
; TITLE: Nucleotide sequence of a functional
; TITLE: cDNA for human thymidylate synthase
; JOURNAL: Nucleic Acids Res.
; VOLUME: 13
; ISSUE: 6
; PAGES: 2035-2043
; DATE: 25-MAR-1985

PCT-US99-12080-1

Db 1321 CAAAACATGATGTGCAATTTCAATCCACAGTACTTATAAAGAAGTGTGGTGAATTTCCAC 1380
Qy 1381 AAGCTATTTTGGAAATATTTTGAATATTTTAAGAAATTTTCAAGCTATTCCTTCAAT 1440
Db 1381 AAGCTATTTTGGAAATATTTTGAATATTTTGAATATTTTGAAGATTTTCAAGCTATTCCTTCAAT 1440
Qy 1441 CTGAGGGAGCTGAGTAACACCATCGATCATGATGTAGAGTGTGGTATGAACTTTTANAGT 1500
Db 1441 CTGAGGGAGCTGAGTAACACCATCGATCATGATGTAGAGTGTGGTATGAACTTTTANAGT 1500
Qy 1501 TGTATTATATGTTGCTATATAAAGAAGTGTCTGC 1536
Db 1501 TGTATTATATGTTGCTATATAAAGAAGTGTCTGC 1536

RESULT 4
US-09-442-384B-792
; Sequence 792, Application US/09442384B
; GENERAL INFORMATION:
; APPLICANT: Chenchik, Alex
; APPLICANT: Lukashov, Matvey
; TITLE OF INVENTION: Hematology/Immunology Array
; FILE REFERENCE: CLON-006CIP15
; CURRENT APPLICATION NUMBER: US/09/442,384B
; CURRENT FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: 09/053,375
; PRIOR FILING DATE: 1998-03-31
; NUMBER OF SEQ ID NOS: 830
; SOFTWARE: FastSeq For Windows Version 4.0
; SEQ ID NO 792
; LENGTH: 1536
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-442-384B-792

Query Match 99.8%; Score 1533; DB 18; Length 1536;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1533; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGACCACTTGGCTGCTCCGTCGGCCGGCCACCTTGGCTGCTCCCT 60
Db 1 GGGGGGGGGGGGACCACTTGGCTGCTCCGTCGGCCGGCCACCTTGGCTGCTCCCT 60
Qy 61 CCGCGCGCGGCACTTGGCTGCTCCGTCGGCCGGCCACCTTGGCTGCTCCCT 120
Db 61 CCGCGCGCGGCACTTGGCTGCTCCGTCGGCCGGCCACCTTGGCTGCTCCCT 120
Qy 121 TCGGAGCTGCGCGCGGCGCTTGCCTCCGCGCGGCGGCGGCGGCGGCGGCGG 180
Db 121 TCGGAGCTGCGCGCGGCGCTTGCCTCCGCGCGGCGGCGGCGGCGGCGGCGG 180
Qy 181 CCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
Db 181 CCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
Qy 241 AGGAAGGAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
Db 241 AGGAAGGAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
Qy 301 AGCTGAGAGATGAAATTCCTTGTGATGAGGAGGAGGAGGAGGAGGAGGAGG 360
Db 301 AGCTGAGAGATGAAATTCCTTGTGATGAGGAGGAGGAGGAGGAGGAGGAGG 360
Qy 361 GAGGAGTGTGTTGTTTATCAAGGAGTCCAAATGCTTAAAGAGTGTCTTCAAGGGA 420
Db 361 GAGGAGTGTGTTGTTTATCAAGGAGTCCAAATGCTTAAAGAGTGTCTTCAAGGGA 420
Qy 421 GTGAAATCTGGGATGCGAATGGATCCCGAGACTTTTGGAGAGCTTGGATTTCCACC 480
Db 421 GTGAAATCTGGGATGCGAATGGATCCCGAGACTTTTGGAGAGCTTGGATTTCCACC 480
Qy 481 AGAGAAGAGGGGACTTGGGCGGCGGCTTATGGCTTCCAGTGGAGGATTTTGGGCGGAGAA 540

```

; Sequence 141, Application US/09631275
; GENERAL INFORMATION:
; APPLICANT: Hong Chen
; APPLICANT: Nelson B. Freimer
; APPLICANT: Thomas Novak
; TITLE OF INVENTION: Method and Compositions for Diagnosing and Treating Chromosome-18
; FILE REFERENCE: 7853-146-999
; CURRENT APPLICATION NUMBER: US/09/631,275
; CURRENT FILING DATE: 2000-08-02
; PRIOR APPLICATION NUMBER: 09/268,992
; PRIOR FILING DATE: 1999-03-16
; PRIOR APPLICATION NUMBER: 09/236,134
; PRIOR FILING DATE: 1999-01-22
; PRIOR APPLICATION NUMBER: 60/078,044
; PRIOR FILING DATE: 1998-03-16
; PRIOR APPLICATION NUMBER: 60/088,312
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/106,056
; PRIOR FILING DATE: 1998-10-28
; NUMBER OF SEQ ID NOS: 158
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 141
; LENGTH: 1536
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-631-275-141

Query Match          99.8%; Score 1533; DB 24; Length 1536;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1533; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGGGGGGGGGGGACCACTTGGCTGCTCCGTCGCCGCGGCACTTGGCTGCTCCGT 60
DB 1 GGGGGGGGGGGGACCACTTGGCTGCTCCGTCGCCGCGGCACTTGGCTGCTCCGT 60
QY 61 CCCCCTCCGCACTTGGCTGCTCCGTCGCCGCGGCACTTGGCTGCTCCGT 120
DB 61 CCCCCTCCGCACTTGGCTGCTCCGTCGCCGCGGCACTTGGCTGCTCCGT 120
QY 121 TGGAGAGTGCCTGGCTGCTCCGTCGCCGCGGCACTTGGCTGCTCCGT 180
DB 121 TGGAGAGTGCCTGGCTGCTCCGTCGCCGCGGCACTTGGCTGCTCCGT 180
QY 181 CCGCGCGGCACTTGGCTGCTCCGTCGCCGCGGCACTTGGCTGCTCCGT 240
DB 181 CCGCGCGGCACTTGGCTGCTCCGTCGCCGCGGCACTTGGCTGCTCCGT 240
QY 241 AGAAGAGGACGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGAC 300
DB 241 AGAAGAGGACGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGAC 300
QY 301 AGCTGAGAGTGAATTCCTCTGCTGACCAACAAAGTGTCTGGAAGGTTTGG 360
DB 301 AGCTGAGAGTGAATTCCTCTGCTGACCAACAAAGTGTCTGGAAGGTTTGG 360
QY 361 GAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
DB 361 GAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
QY 421 GTGAAATCTGGGATCCCAATGATCCGAGACTTTTGGACAGCTTCCACCC 480
DB 421 GTGAAATCTGGGATCCCAATGATCCGAGACTTTTGGACAGCTTCCACCC 480
QY 481 AGAGAGAGGAGGACTTGGGACCACTTATGCTTCCAGTGAAGGCAATTTGGGACAAA 540
DB 481 AGAGAGAGGAGGACTTGGGACCACTTATGCTTCCAGTGAAGGCAATTTGGGACAAA 540
QY 541 TACAGAGATATGAGATCAATTAATTCAGAGAGAGGAGGAGGAGGAGGAGGAGGAGG 600
DB 541 TACAGAGATATGAGATCAATTAATTCAGAGAGAGGAGGAGGAGGAGGAGGAGGAGG 600
QY 601 GACACGATCAAAACGACCTTACGACGAGAGATATATATGTCGCTTGGATCCAGA 660

```

```

DB 601 GACACGATCAAAACGACCTTACGACGAGAGATCAATGATGCTTGGATCCAGA 660
QY 661 GATCTCTCTGATAGGCGCTGCTCCATGCAATGCTCCGCTGCTGCTGCTGCTGCTGCT 720
DB 661 GATCTCTCTGATAGGCGCTGCTCCATGCAATGCTCCGCTGCTGCTGCTGCTGCTGCT 720
QY 721 AGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
DB 721 AGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
QY 781 AAGATGCGGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
DB 781 AAGATGCGGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
QY 841 GGTGACTTATACACTTGGAGATGACATATTAATTAATTAATTAATTAATTAATTAATTA 900
DB 841 GGTGACTTATACACTTGGAGATGACATATTAATTAATTAATTAATTAATTAATTAATTA 900
QY 901 AAAATTCAGCTTACGAGAACCCAGACTTTCGCAAGCTGAGATTTCCGAAAAGTT 960
DB 901 AAAATTCAGCTTACGAGAACCCAGACTTTCGCAAGCTGAGATTTCCGAAAAGTT 960
QY 961 GAGAAATTTGATGACTTCAAGCTGAAGACTTTCGCAAGCTGAGATTTCCGAAAAGTT 1020
DB 961 GAGAAATTTGATGACTTCAAGCTGAAGACTTTCGCAAGCTGAGATTTCCGAAAAGTT 1020
QY 1021 ACTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1080
DB 1021 ACTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1080
QY 1081 GCTCTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
DB 1081 GCTCTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
QY 1141 AACTAGGTCAAAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
DB 1141 AACTAGGTCAAAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
QY 1201 GCGAAATGTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
DB 1201 GCGAAATGTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
QY 1261 GATCTGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
DB 1261 GATCTGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
QY 1321 CAATAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380
DB 1321 CAATAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380
QY 1381 AAGCTATTTTGGAAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGA 1440
DB 1381 AAGCTATTTTGGAAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGA 1440
QY 1441 CTGAGAGGCTGAGTAAACATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
DB 1441 CTGAGAGGCTGAGTAAACATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
QY 1501 TGTATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560
DB 1501 TGTATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560

RESULT 6
US-09-658-659-6
; Sequence 6, Application US/09658659
; GENERAL INFORMATION:
; APPLICANT: Stanton, Jr., Vincent P.
; TITLE OF INVENTION: GENB SEQUENCE VARIANCE IN GENES RELATED
; TITLE OF INVENTION: TO POLAR METABOLISM HAVING UTILITY IN DETERMINING THE
; FILE REFERENCE: 11926-015001

```

```

; CURRENT APPLICATION NUMBER: US/09/658,659
; CURRENT FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 09/596,033
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 09/357,743
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 09/357,024
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: 60/093,484
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1536
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1066
; OTHER INFORMATION: n = t or c
; NAME/KEY: misc_feature
; LOCATION: 1136
; OTHER INFORMATION: n = a or g
; NAME/KEY: misc_feature
; LOCATION: 1497
; OTHER INFORMATION: n = t or a
; US-09-658-659-6

```

```

Query Match          99.8%; Score 1533; DB 25; Length 1536;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1536; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 GGGGGGGGGGACACTTGGCTGCTCCGCGCGGCACTTGGCTGCTCCGCT 60
DB 1 GGGGGGGGGGAGCACTTGGCTGCTCCGCGCGGCACTTGGCTGCTCCGCT 60
QY 61 CCGCGCGCGCACTTGGCTGCTCCGCGCGGCACTTGGCTGCTCCGCT 120
DB 61 CCGCGCGCGCACTTGGCTGCTCCGCGCGGCACTTGGCTGCTCCGCT 120
QY 121 TCGGAGCTGCGCGCGCGCTTGGCTGCTCCGCGCGGCACTTGGCTGCTCCGCT 180
DB 121 TCGGAGCTGCGCGCGCGCTTGGCTGCTCCGCGCGGCACTTGGCTGCTCCGCT 180
QY 181 CCGCGCGCGAGCTGCACTTGGCTGCTCCGCGCGGCACTTGGCTGCTCCGCT 240
DB 181 CCGCGCGCGAGCTGCACTTGGCTGCTCCGCGCGGCACTTGGCTGCTCCGCT 240
QY 241 AGAAGGACGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGAC 300
DB 241 AGAAGGACGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCGGAC 300
QY 301 AGCTTGAGATGATTCCTCTGCTGACCAACCTGCTGCTGCTGCTGCTGCTGCT 360
DB 301 AGCTTGAGATGATTCCTCTGCTGACCAACCTGCTGCTGCTGCTGCTGCTGCT 360
QY 361 GAGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
DB 361 GAGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
QY 421 GTGAAAATCTGGGATGCCAATGATCCCGAAGCTTTTGGACAGGCTGGGATTTCTCCACC 480
DB 421 GTGAAAATCTGGGATGCCAATGATCCCGAAGCTTTTGGACAGGCTGGGATTTCTCCACC 480
QY 481 AGAAGAAAGGGGACTTGGGCGGCTTTATGCTTCAATGGAAGGCAATTTTGGGGGAGAA 540
DB 481 AGAAGAAAGGGGACTTGGGCGGCTTTATGCTTCAATGGAAGGCAATTTTGGGGGAGAA 540
QY 541 TACAGATATGGAATCAGATTTATTCAGAGCAGAGGCTGACCACTGCAAAAGAGTATTT 600
DB 541 TACAGATATGGAATCAGATTTATTCAGAGCAGAGGCTGACCACTGCAAAAGAGTATTT 600
QY 601 GACACCATCAAAACCACTGACGACAGAGAAATCANTGAGCTTGGGAAATCCAGAA 660

```

```

DB 601 GACACCATCAAAACCACTGACGACAGAGAAATCANTGAGCTTGGGAAATCCAGAA 660
QY 661 GATCTTCCTGATGAGGCGCTGCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
DB 661 GATCTTCCTGATGAGGCGCTGCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
QY 721 AGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
DB 721 AGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
QY 781 AACATGCGACGCTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
DB 781 AACATGCGACGCTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
QY 841 GGTGACTTTATACACTTGGGAGATGCAATTTTACTGATGACATCGAGCCACTG 900
DB 841 GGTGACTTTATACACTTGGGAGATGCAATTTTACTGATGACATCGAGCCACTG 900
QY 901 AAAATTCAGCTTCAGGAGAACCCAGACCTTCCGAAAGCTCAGGATTCGGAAGATT 960
DB 901 AAAATTCAGCTTCAGGAGAACCCAGACCTTCCGAAAGCTCAGGATTCGGAAGATT 960
QY 961 GAGAAATTTGATGACTTCAAGCTGAGACTTCAATGGAAGGTAACAATCCGATCCA 1020
DB 961 GAGAAATTTGATGACTTCAAGCTGAGACTTCAATGGAAGGTAACAATCCGATCCA 1020
QY 1021 ACTATTAATGAAATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
DB 1021 ACTATTAATGAAATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
QY 1081 GTCTTTAGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
DB 1081 GTCTTTAGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
QY 1141 AACTAGGTCAAAATTCGTCGCGGACCTATCATGATTAATTTTAAAGATGTCGACT 1200
DB 1141 AACTAGGTCAAAATTCGTCGCGGACCTATCATGATTAATTTTAAAGATGTCGACT 1200
QY 1201 GGCAAATGTAATCTGCGCAATTCCTTCCATTAATAAAGGCTTTGAGTTACTCACTGAG 1260
DB 1201 GGCAAATGTAATCTGCGCAATTCCTTCCATTAATAAAGGCTTTGAGTTACTCACTGAG 1260
QY 1261 GTATCTGCAATGCTGAGTTATGAAACAAGTGAAGAAATGATGCTCTTAA 1320
DB 1261 GTATCTGCAATGCTGAGTTATGAAACAAGTGAAGAAATGATGCTCTTAA 1320
QY 1321 CAAAACATGATGATGATTTCAATCCACGTAATTAAGAAGTGGGAAATTCAC 1380
DB 1321 CAAAACATGATGATGATTTCAATCCACGTAATTAAGAAGTGGGAAATTCAC 1380
QY 1381 AAGCTATTTTGAATATTTTGAATATTTTGAATATTTTGAATATTTTGAATATTTT 1440
DB 1381 AAGCTATTTTGAATATTTTGAATATTTTGAATATTTTGAATATTTTGAATATTTT 1440
QY 1441 CTGAGGAGCTGAGTAAACCAATTCATCATGATGATGATGATGATGATGATGAT 1500
DB 1441 CTGAGGAGCTGAGTAAACCAATTCATCATGATGATGATGATGATGATGATGAT 1500
QY 1501 TGTTTTATGCTGCTATATTAAGAAGTCTTCTGC 1536
DB 1501 TGTTTTATGCTGCTATATTAAGAAGTCTTCTGC 1536

```

```

RESULT 7
US-09-877-178-11
; Sequence 11, Application US/09877178
; GENERAL INFORMATION:
; APPLICANT: Kathleen Danenberg, et al.
; TITLE OF INVENTION: METHOD OF DETERMINING A CHEMOTHERAPEUTIC
; TITLE OF INVENTION: REGIMEN BASED ON ERCC1 and TS EXPRESSION
; FILE REFERENCE: 11220-119
; CURRENT APPLICATION NUMBER: US/09/877,178

```


CURRENT FILING DATE: 2001-06-11
 NUMBER OF SEQ ID NOS: 11
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 11
 LENGTH: 1536
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-09-877-178-11

Query Match 99.8%; Score 1533; DB 33; Length 1536;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1533; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

1 GGGGGGGGGGGGACACTTGGCTGCTCCGCTCCGCGCGCACTTGGCTGCTCCGCT
 1 GGGGGGGGGGGGACACTTGGCTGCTCCGCTCCGCGCGCACTTGGCTGCTCCGCT
 61 CCGCGCGCGCACTTGGCTGCTCCGCTCCGCGCGCGCGCGCGCGCGCGCGCGCG
 61 CCGCGCGCGCACTTGGCTGCTCCGCTCCGCGCGCGCGCGCGCGCGCGCGCGCG
 121 TCGGAGCTGCG
 121 TCGGAGCTGCG
 181 CCG
 181 CCG
 241 AGGAGGAGGAGCG
 241 AGGAGGAGGAGCG
 301 AGCTGAGAGTGAATTCCTCTGCTGACAAACCAACGCTGCTTGGAGAGGCTTTG
 301 AGCTGAGAGTGAATTCCTCTGCTGACAAACCAACGCTGCTTGGAGAGGCTTTG
 361 GAGGAGTGTGCTGCTGCTTATCAAGGAGTCCAAATGCTTAAAGAGTGTCTTCAAGG
 361 GAGGAGTGTGCTGCTGCTTATCAAGGAGTCCAAATGCTTAAAGAGTGTCTTCAAGG
 421 GTGAAAATCTGGAGATCCCAATGATCCCGACATCTTTGGACAGCTTGGAGATCCCA
 421 GTGAAAATCTGGAGATCCCAATGATCCCGACATCTTTGGACAGCTTGGAGATCCCA
 481 AGAAGAAAGAGGAGTGTGGGCGCAATTTATGCTTCCAGTGGAGGATTTTGGAGAA
 481 AGAAGAAAGAGGAGTGTGGGCGCAATTTATGCTTCCAGTGGAGGATTTTGGAGAA
 541 TACAGAGTATGATCAATTAATTCAGAGCAAGGAGTTCACACTGCAAAAGTGAAT
 541 TACAGAGTATGATCAATTAATTCAGAGCAAGGAGTTCACACTGCAAAAGTGAAT
 601 GACACATCAAAAACCAACCTTGAAGCAAGAGAAATCAATGCTGCTTGGAGATCCAGA
 601 GACACATCAAAAACCAACCTTGAAGCAAGAGAAATCAATGCTGCTTGGAGATCCAGA
 661 GATCTTCTCTGATGCGCTGCTCCATGCAATGCTTCCAGTTCATGCTGAGAC
 661 GATCTTCTCTGATGCGCTGCTCCATGCAATGCTTCCAGTTCATGCTGAGAC
 721 AGTGAAGCTGCTTCCAGCTGCTTACAGAGATCGGAGAGCAATGCTGCTGCTTTC
 721 AGTGAAGCTGCTTCCAGCTGCTTACAGAGATCGGAGAGCAATGCTGCTGCTTTC
 781 AACATCGCGAGCTAGCGCTGCTCAGATCAATGCTGCTGCTGCTGCTGCTGCTGCT
 781 AACATCGCGAGCTAGCGCTGCTCAGATCAATGCTGCTGCTGCTGCTGCTGCTGCT
 841 GGTGACTTATACACTTTGGAGAGTGAACATATTTTCTGATGACATGAGCACTG
 841 GGTGACTTATACACTTTGGAGAGTGAACATATTTTCTGATGACATGAGCACTG

901 AAAATTCAGCTTCAGGAGAACCCAGACCTTCCCAAGCTCAGAGTCTTGAAAAGTT
 901 AAAATTCAGCTTCAGGAGAACCCAGACCTTCCCAAGCTCAGAGTCTTGAAAAGTT
 961 GAGAAAATGATGACTTCAAGCTGAAGACTTCAAGTGAAGGGTCAATCCGCTCA
 961 GAGAAAATGATGACTTCAAGCTGAAGACTTCAAGTGAAGGGTCAATCCGCTCA
 1021 ACTATTAATGAAATGCTGCTTGAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCA
 1021 ACTATTAATGAAATGCTGCTTGAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCA
 1081 GTCCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGA
 1081 GTCCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGA
 1141 AACTAGGTCAAAAATCTGCTCCGCTGACCTATCAATTAATTTAAGAGTGTGCACT
 1141 AACTAGGTCAAAAATCTGCTCCGCTGACCTATCAATTAATTTAAGAGTGTGCACT
 1201 GCGAAATGTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT
 1201 GCGAAATGTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT
 1261 GTATCTGACATCTGAGGCTTATGAACAAAGTGAAGGAGTGAAGTGAAGTGAAGTGA
 1261 GTATCTGACATCTGAGGCTTATGAACAAAGTGAAGGAGTGAAGTGAAGTGAAGTGA
 1321 CAAAACATGATGATGATTCATTCATCCAGTACTTATTAAGAGTGTGCACT
 1321 CAAAACATGATGATGATTCATTCATCCAGTACTTATTAAGAGTGTGCACT
 1381 AAGCTATTTTGAATATTTTGAATATTTTGAATATTTTGAATATTTTGAATATTTTGA
 1381 AAGCTATTTTGAATATTTTGAATATTTTGAATATTTTGAATATTTTGAATATTTTGA
 1441 CTGAGGAGCTGAGTACCATGATGATGATGATGATGATGATGATGATGATGATGAT
 1441 CTGAGGAGCTGAGTACCATGATGATGATGATGATGATGATGATGATGATGATGAT
 1501 TGTATTATATGCTATATATTAAGAGTGTCTG
 1501 TGTATTATATGCTATATATTAAGAGTGTCTG

RESULT 8
 US-09-962-665-6
 Sequence 6, Application US/09962665
 GENERAL INFORMATION:
 APPLICANT: Stanton, Jr., Vincent P.
 TITLE OF INVENTION: POLYMERIZABLE SYNTHETASE GENE SEQUENCE
 TITLE OF INVENTION: VARIANCES HAVING UTILITY IN DETERMINING THE
 TITLE OF INVENTION: TREATMENT OF DISEASE
 FILE REFERENCE: 11926-015004
 CURRENT APPLICATION NUMBER: US/09/962,665
 CURRENT FILING DATE: 2001-09-24
 PRIOR APPLICATION NUMBER: 09/658,659
 PRIOR FILING DATE: 2000-09-08
 PRIOR APPLICATION NUMBER: 09/596,033
 PRIOR FILING DATE: 2000-06-15
 PRIOR APPLICATION NUMBER: 09/357,743
 PRIOR FILING DATE: 1999-07-20
 PRIOR APPLICATION NUMBER: 09/357,024
 PRIOR FILING DATE: 1999-07-19
 PRIOR APPLICATION NUMBER: 60/093,484
 PRIOR FILING DATE: 1998-07-20
 NUMBER OF SEQ ID NOS: 16
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 6
 LENGTH: 1536
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:


```

; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1066
; OTHER INFORMATION: n = t or c
; NAME/KEY: misc_feature
; LOCATION: 1136
; OTHER INFORMATION: n = a or g
; NAME/KEY: misc_feature
; LOCATION: 1497
; OTHER INFORMATION: n = t or a
; US-09-962-677-6

```

```

Query Match          99.8%; Score 1533; DB 36; Length 1536;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1536; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 GGGGGGGGGGGGACCACTTGGCTGCTCCGTCGCGCGCGCCACTTGGCTGCTCCGCT 60
Db 1 GGGGGGGGGGGGACCACTTGGCTGCTCCGTCGCGCGCGCCACTTGGCTGCTCCGCT 60
OY 61 CCCGCGCGCCACTTGGCTGCTCCGTCGCGCGCGCGCGCGCCACTTGGCTGCTCCGCT 120
Db 61 CCCGCGCGCCACTTGGCTGCTCCGTCGCGCGCGCGCGCGCGCCACTTGGCTGCTCCGCT 120
OY 121 TCGGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
Db 121 TCGGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
OY 181 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240
Db 181 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240
OY 241 AGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
Db 241 AGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
OY 301 AGCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
Db 301 AGCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
OY 361 GAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
Db 361 GAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
OY 421 GTGAAATCTGGGATGCGCATGATCCGAGACTTTTGGAGAGGCTTCCAGC 480
Db 421 GTGAAATCTGGGATGCGCATGATCCGAGACTTTTGGAGAGGCTTCCAGC 480
OY 481 AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
Db 481 AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
OY 541 TACAGAGATGAGATGAGATGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 600
Db 541 TACAGAGATGAGATGAGATGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 600
OY 601 GACACATCAAAACCAACCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
Db 601 GACACATCAAAACCAACCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
OY 661 GATCTTCTCTGATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Db 661 GATCTTCTCTGATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
OY 721 AGTGAAGCTGCTGCGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
Db 721 AGTGAAGCTGCTGCGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
OY 781 AACATGCGCAGCTACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Db 781 AACATGCGCAGCTACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840

```

```

OY 841 GGTGACTTTATACACTTTGGGAGATGACATATTTACCTGATTCACATCGAGCCACTG 900
Db 841 GGTGACTTTATACACTTTGGGAGATGACATATTTACCTGATTCACATCGAGCCACTG 900
OY 901 AAAATTCAGCTTCAGCAGAACCCAGACCTTTCCCAAGCTCAGAGATTTCCAAAAGTT 960
Db 901 AAAATTCAGCTTCAGCAGAACCCAGACCTTTCCCAAGCTCAGAGATTTCCAAAAGTT 960
OY 961 GAGAAATTTGATGACTTCAAGCTGAAGACTTTCAGATTGAAGGTTCAATCCGATTCGA 1020
Db 961 GAGAAATTTGATGACTTCAAGCTGAAGACTTTCAGATTGAAGGTTCAATCCGATTCGA 1020
OY 1021 ACTATTAATGAAATGAGCTGTTTGAAGGCTTTCAAGAGCTTGAAGATTTGTGA 1080
Db 1021 ACTATTAATGAAATGAGCTGTTTGAAGGCTTTCAAGAGCTTGAAGATTTGTGA 1080
OY 1081 GTCTTTAGGGGTTGGGCTGAGATGCGAGGTTAAAGTTCTTTTGTCTTAAAGAAAGAG 1140
Db 1081 GTCTTTAGGGGTTGGGCTGAGATGCGAGGTTAAAGTTCTTTTGTCTTAAAGAAAGAG 1140
OY 1141 AACTAGGTCAAAAATCTGCTCGTGAACCTATCAGTTATTAATTTTAAAGATGTTGCAC 1200
Db 1141 AACTAGGTCAAAAATCTGCTCGTGAACCTATCAGTTATTAATTTTAAAGATGTTGCAC 1200
OY 1201 GCGAATGTAACCTGCGAGCTTTTCCATATTAAGAGCTTTGAGTTACTCACTGAG 1260
Db 1201 GCGAATGTAACCTGCGAGCTTTTCCATATTAAGAGCTTTGAGTTACTCACTGAG 1260
OY 1261 GTATCTGACAACTGAGGTTATGACAAAGTGAAGAGATGAAATGATGCTTTAG 1320
Db 1261 GTATCTGACAACTGAGGTTATGACAAAGTGAAGAGATGAAATGATGCTTTAG 1320
OY 1321 CAAAACATGATGTCATTTTCAATCCACGTCATTAATAAGAGAGTTGGATTTTCA 1380
Db 1321 CAAAACATGATGTCATTTTCAATCCACGTCATTAATAAGAGAGTTGGATTTTCA 1380
OY 1381 AAGCTATTTTGGATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAAT 1440
Db 1381 AAGCTATTTTGGATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAAT 1440
OY 1441 CTGAGGAGCTGAGTAAACCATCATGATGATGATGATGATGATGATGATGATGATGAT 1500
Db 1441 CTGAGGAGCTGAGTAAACCATCATGATGATGATGATGATGATGATGATGATGATGAT 1500
OY 1501 TGTTTTATGTTGCTATATTAAGAGATGTTCTGC 1536
Db 1501 TGTTTTATGTTGCTATATTAAGAGATGTTCTGC 1536

```

```

RESULT 10
US-09-963-333-6
; Sequence 6, Application US/0996333
; GENERAL INFORMATION:
; APPLICANT: Stanton, Jr., Vincent P.
; TITLE OF INVENTION: THYMIDINE SYNTHASE GENE SEQUENCE VARIANCES
; TITLE OF INVENTION: HAVING UTILITY IN DETERMINING THE TREATMENT
; FILE REFERENCE: 11926-015002
; CURRENT APPLICATION NUMBER: US/09/963,333
; PRIOR APPLICATION NUMBER: 2001-09-24
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 09/596,033
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 09/357,743
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 09/357,024
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: 60/093,484
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6

```

Db	781	AAATGCGACGTCAGCCCTGCTCAACATATATGACGACATCAAGGGCCATGAAGCCA	840
Oy	841	GGTACTTTATACACACTTTGGAGATAGACATATTTACCGAATCACTTCAGAGCACTG	9000
Db	841	GGGACTTTATACACACTTTGGAGATAGACATATTTACTCGAATTCACATCGAGCACTG	9000
Oy	901	AAATATCGCTTAGCGGAGAACCCAGACTTTCCCAAGCTCAGSATTCTTGAAAAGTT	9600
Db	901	AAATATCGCTTAGCGGAGAACCCAGACTTTCCCAAGSCTCAAGATCTTGAAAAGTT	9600
Oy	961	GAGAAAATGTATGACTTCAAACTGATAGACTTCAGATTGAAGGGTCAATCCGATCCA	10200
Db	961	GAGAAAATGTATGACTTCAAACTGATAGACTTCAGATTGAAGGGTCAATCCGATCCA	10200
Oy	1021	ACATTTAAAAAGAAATGCGCTGTTTAGGGTGCCTTCAAGSAGTTGAAAGCATATTGCA	10800
Db	1021	ACATTTAAAAAGAAATGCGCTGTTTAGGGTGCCTTCAAGSAGTTGAAAGCATATTGCA	10800

Db	1021	ACTATTAAAAATGGAATATGCTGTTTATAGGATGCTTTCAAAAGAGCTTGAAGATATTGTCA	1081
QY	1081	GTCTTTAGGAGGTGAGCAATGCGATGAGCCGAGATAAAGTCTTTTGTCTCTAAAGANNAAG	114
Db	1081	GTTCTTTAGGAGGTGAGGCTGATGCGAGGATAAAGTCTTTTGTCTCTAAAGANNAAG	114
QY	1141	AACATGTCGCAAAAATCTGTCGCGTACCTATCACTTAATTTTAAGATGTTGCCACT	120
Db	1141	AACATGTCGCAAAAATCTGTCGCGTACCTATCACTTAATTTTAAGATGTTGCCACT	120
QY	1201	GGAATAATGTAATCTGTCCAGTCTTTCCTCAATAAAGGCTTGAAGTATCACTAGAG	126
Db	1201	GGAATAATGTAATCTGTCCAGTCTTTCCTCAATAAAGGCTTGAAGTATCACTAGAG	126
QY	1261	GTATCTGACATGCTGAGGTTATGACAAAGTGAAGAGATGAATATATGTCCTTAG	132
Db	1261	GTATCTGACATGCTGAGGTTATGACAAAGTGAAGAGATGAATATATGTCCTTAG	132
QY	1321	CAAAAACATGATGTCGATTTGATCCCGATCTTATTAAGAGGTTGCTGAATTTGAC	138
Db	1321	CAAAAACATGATGTCGATTTGATCCCGATCTTATTAAGAGGTTGCTGAATTTGAC	138
QY	1381	AAGCTATTTTGGAAATATTTTAAATTTTAAGATTTTACACAGCTATTCCTCAAT	144
Db	1381	AAGCTATTTTGGAAATATTTTAAATTTTAAAGATTTTACACAGTATTCCTCAAT	144
QY	1441	CTGAGGAGCTGAGTAAACACATCGATCATGATGTGAGTGTGGTTATGAATTTANAGT	150
Db	1441	CTGAGGAGCTGAGTAAACACATCGATCATGATGTGAGTGTGGTTATGAATTTANAGT	150
QY	1501	TGTTTATATGTTGCTATATTAATAAGAGTGTCTGC	1536
Db	1501	TGTTTATATGTTGCTATATTAATAAGAGTGTCTGC	1536
RESULT 11			
US-10-172-118-556			
Sequence 556, Application US/10172118			
GENERAL INFORMATION:			
APPLICANT: Dai, Hongyue			
APPLICANT: He, Yundong			
APPLICANT: Linsley, Peter			
APPLICANT: Mao, Mao			
APPLICANT: Roberts, Chris			
APPLICANT: Van 't Veer, Laura			
APPLICANT: Van de Vijver, Marc			
APPLICANT: Bernards, Rene			
TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients			
FILE REFERENCE: 9301-175-999			
CURRENT APPLICATION NUMBER: US/10-172-118			
CURRENT FILING DATE: 2002-06-14			
PRIOR APPLICATION NUMBER: 60/380, 770			
PRIOR FILING DATE: 2002-05-14			
NUMBER OF SEQ ID NOS: 2699			
SEQ ID NO 556			
LENGTH: 1536			
TYPE: DNA			

ORGANISM: Homo sapiens
 PUBLICATION INFORMATION:
 DATABASE ACCESSION NUMBER: NM_001071
 DATABASE ENTRY DATE: 2001-06-18
 US-10-172-118-556

Query Match 99.8%; Score 1533; DB 41; Length 1536;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1533; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

1 GGGGGGGGGGGGACCTTGGCTGCTCCGCTCCCGCGCGACCTTGGCTGCTCCGCT 60
 1 GGGGGGGGGGGGACCTTGGCTGCTCCGCTCCCGCGCGACCTTGGCTGCTCCGCT 60
 61 CCGCGCGCGACCTTGGCTGCTCCGCTCCCGCGCGACCTTGGCTGCTCCGCT 120
 61 CCGCGCGCGACCTTGGCTGCTCCGCTCCCGCGCGACCTTGGCTGCTCCGCT 120
 121 TCGGAGCTGCGCGCGCGCTTGGCTCCCGCGCGACCTTGGCTGCTCCGCT 180
 121 TCGGAGCTGCGCGCGCGCTTGGCTCCCGCGCGACCTTGGCTGCTCCGCT 180
 181 CCGCGCGCGACCTTGGCTGCTCCGCTCCCGCGCGACCTTGGCTGCTCCGCT 240
 181 CCGCGCGCGACCTTGGCTGCTCCGCTCCCGCGCGACCTTGGCTGCTCCGCT 240
 241 AGGAG 300
 241 AGGAG 300
 301 AGGAG 360
 301 AGGAG 360
 361 GAG 420
 361 GAG 420
 421 GTGAG 480
 421 GTGAG 480
 481 AGGAG 540
 481 AGGAG 540
 541 TACGAG 600
 541 TACGAG 600
 601 GAG 660
 601 GAG 660
 661 GATCT 720
 661 GATCT 720
 721 AGTGAAGCTGCTGCGAGCTGTAACAAGAGAGAGAGAGAGAGAGAGAGAGAG 780
 721 AGTGAAGCTGCTGCGAGCTGTAACAAGAGAGAGAGAGAGAGAGAGAGAGAG 780
 781 AACATGCGAG 840
 781 AACATGCGAG 840
 841 GGTGAGCTTATACACACTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
 841 GGTGAGCTTATACACACTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
 901 AAAATTGAGCTTACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
 901 AAAATTGAGCTTACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960

961 GAG 1020
 961 GAG 1020
 1021 ACTATTTAAATGAG 1080
 1021 ACTATTTAAATGAG 1080
 1081 GTCTTTAG 1140
 1081 GTCTTTAG 1140
 1141 AACTAG 1200
 1141 AACTAG 1200
 1201 GGCAG 1260
 1201 GGCAG 1260
 1261 GTATCTGACATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
 1261 GTATCTGACATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
 1321 CAAAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
 1321 CAAAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
 1381 AAGCTATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGA 1440
 1381 AAGCTATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGA 1440
 1441 CTGAG 1500
 1441 CTGAG 1500
 1501 TGTATTTATGTTGCTATTAATAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1536
 1501 TGTATTTATGTTGCTATTAATAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1536

RESULT 12
 US-60-412-049-350
 Sequence 350, Application US/60412049
 GENERAL INFORMATION:
 APPLICANT: Baker, Joffe B.
 APPLICANT: Cronin, Maureen T.
 APPLICANT: Kiefer, Michael C.
 APPLICANT: Shak, Steve
 APPLICANT: Walker, Michael Graham
 TITLE OF INVENTION: GENE EXPRESSION PROFILING IN BIOESIED TUMOR TISSUES
 FILE REFERENCE: GANHE.004PR
 CURRENT APPLICATION NUMBER: US/60/412,049
 CURRENT FILING DATE: 2002-09-18
 NUMBER OF SEQ ID NOS: 384
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 350
 LENGTH: 1536
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-60-412-049-350

Query Match 99.8%; Score 1533; DB 85; Length 1536;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1533; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

1 GGGGGGGGGGGGACCTTGGCTGCTCCGCTCCCGCGCGACCTTGGCTGCTCCGCT 60
 1 GGGGGGGGGGGGACCTTGGCTGCTCCGCTCCCGCGCGACCTTGGCTGCTCCGCT 60
 61 CCGCGCGCGACCTTGGCTGCTCCGCTCCCGCGCGACCTTGGCTGCTCCGCT 120
 61 CCGCGCGCGACCTTGGCTGCTCCGCTCCCGCGCGACCTTGGCTGCTCCGCT 120

Db 213 GGGGGGGGGGGGACCACTTGGCTGCTCCGCTCCGCGCGCACCTTGGCTCCGCT 272
 Qy 61 CCCCGCGGCACTTGGCTGCTCCGCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
 Db 273 CCCCGCGGCACTTGGCTGCTCCGCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 332
 Qy 121 TCGGAGCTGCG 180
 Db 333 TCGGAGCTGCG 392
 Qy 181 CCG 240
 Db 393 CCG 452
 Qy 241 AGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300
 Db 453 AGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 512
 Qy 301 AGCCTGAGAGATGAATTCCTGCTGCAACAACGCTGCTGCAACAACGCTGCTGCA 360
 Db 513 AGCCTGAGAGATGAATTCCTGCTGCAACAACGCTGCTGCAACAACGCTGCTGCA 572
 Qy 361 GAGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
 Db 573 GAGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 632
 Qy 421 GTGAAATCTGGGATGCGCAATGATCCCGAGACTTTTGGAGAGCACTTTTGGGAGAA 480
 Db 633 GTGAAATCTGGGATGCGCAATGATCCCGAGACTTTTGGAGAGCACTTTTGGGAGAA 692
 Qy 481 AG 540
 Db 693 AG 752
 Qy 541 TACAGAGATATGAGATCAATTAATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
 Db 753 TACAGAGATATGAGATCAATTAATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 812
 Qy 601 GACACCATCAAAACCAACCTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
 Db 813 GACACCATCAAAACCAACCTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 872
 Qy 661 GATCTTCTCTGAG 720
 Db 873 GATCTTCTCTGAG 932
 Qy 721 AGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
 Db 933 AGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 992
 Qy 781 AACATGCGAGCTAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
 Db 993 AACATGCGAGCTAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1052
 Qy 841 GGTGACTTTATACACATTTGGAGATGACATATTTTCTGATACATGAGCTGAGCTG 900
 Db 1053 GGTGACTTTATACACATTTGGAGATGACATATTTTCTGATACATGAGCTGAGCTG 1112
 Qy 901 AAAATTCAGCTTACGCGAGAACCAAGCTTTTCCAAAGCTAGAGATTTCTTGAAGATT 960
 Db 1113 AAAATTCAGCTTACGCGAGAACCAAGCTTTTCCAAAGCTAGAGATTTCTTGAAGATT 1172
 Qy 961 GAGAAATTTGATGATTTCAAAAGCTGAGAGCTTTCAATTTGAGAGGATTAATTCGACAT 1020
 Db 1173 GAGAAATTTGATGATTTCAAAAGCTGAGAGCTTTCAATTTGAGAGGATTAATTCGACAT 1232
 Qy 1021 ACTATTTAAATGAAATGAGCTGTTTAAAGGCTTTCAAAAGGAGCTGAAAGGATTTGCTCA 1080
 Db 1233 ACTATTTAAATGAAATGAGCTGTTTAAAGGCTTTCAAAAGGAGCTGAAAGGATTTGCTCA 1292
 Qy 1081 GTCTTTTAAAGGCTTGAATGCGGAGGATTAAGTTCTTTTGTCTTAAAGGAGAGAG 1140
 Db 1293 GTCTTTTAAAGGCTTGAATGCGGAGGATTAAGTTCTTTTGTCTTAAAGGAGAGAG 1352

Qy 1141 AACTAGTTCAAAATCTGTCGCTGACCTTACAGTTATTTTAAAGATGTTGCACT 1200
 Db 1353 AACTAGTTCAAAATCTGTCGCTGACCTTACAGTTATTTTAAAGATGTTGCACT 1412
 Qy 1201 GGGCAATTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
 Db 1413 GGGCAATTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1472
 Qy 1261 GTATCTGCAATGCTGAGTTATGAAACAAGTGAAGATGAATGATGCTGCTGCTGCTGCT 1320
 Db 1473 GTATCTGCAATGCTGAGTTATGAAACAAGTGAAGATGAATGATGCTGCTGCTGCTGCT 1532
 Qy 1321 CAAAACATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380
 Db 1533 CAAAACATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1592
 Qy 1381 AAGCTATTTTGGAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAAT 1440
 Db 1593 AAGCTATTTTGGAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAAT 1652
 Qy 1441 CTGAGGAGCTGAGTACACATGATGATGATGATGATGATGATGATGATGATGATGAT 1494
 Db 1653 CTGAGGAGCTGAGTACACATGATGATGATGATGATGATGATGATGATGATGATGAT 1712
 Qy 1495 TAAAGTTGTTTATATGTTGCTAATAATAAAGTTGCTGCTGCTGCTGCTGCTGCTGCT 1536
 Db 1713 TAAAGTTGTTTATATGTTGCTAATAATAAAGTTGCTGCTGCTGCTGCTGCTGCTGCT 1754

RESULT 14
 US-09-652-109-9676
 ; Sequence 9676, Application US/09652109
 ; GENERAL INFORMATION:
 ; APPLICANT: McCarthy, Sean A.
 ; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
 ; FILE REFERENCE: 1600.1180-001
 ; CURRENT APPLICATION NUMBER: US/09/652,109
 ; PRIOR FILING DATE: 2000-08-30
 ; PRIOR APPLICATION NUMBER: 60/151,128
 ; NUMBER OF SEQ ID NOS: 10105
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 9676
 ; LENGTH: 1829
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-652-109-9676

Query Match 98.0%; Score 1505.4; DB 25; Length 1829;
 Best Local Similarity 99.3%; Pred. No. 0;
 Matches 1522; Conservative 0; Mismatches 4; Indels 6; Gaps 1;

Qy 11 GAGCACTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 70
 Db 45 GAGCACTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 104
 Qy 71 CACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 130
 Db 105 CACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 164
 Qy 131 GCGCGCGGCTTGGCG 190
 Db 165 GCGCGCGGCTTGGCG 224
 Qy 191 GGGAGCTGAGTACTGGGGGAGATCCAAACATCTTCCGCTGCGGCTGAGAGAGAGAG 250
 Db 225 GGGAGCTGAGTACTGGGGGAGATCCAAACATCTTCCGCTGCGGCTGAGAGAGAGAG 284
 Qy 251 ACCGAGCGGCGACCGGACCTCTGTGGATTTGCGAGATGCGAGGCGGCTTACAGCTGAGAG 310
 Db 285 ACCGAGCGGCGACCGGACCTCTGTGGATTTGCGAGATGCGAGGCGGCTTACAGCTGAGAG 344

311 ATGATTCCTCTGCTGACAAACCAACGAGTGTTCGGAAGGGTGTTCGAGAGTTC 370
345 ATGAAATCCCTCTGCTGACAAACCAACGAGTGTTCGGAAGGGTGTTCGAGAGTTC 404
371 TGTGTTTATCAAGGATTCACAAATGCTAAAGAGCTCTTCCAGAGGAGTAAATCT 430
405 TGTGTTTATCAAGGATTCACAAATGCTAAAGAGCTCTTCCAGAGGAGTAAATCT 464
431 GGGATGCCAATGATCCCGAGACTTTTGGACAGCTGGGATCTCCACAGAGAGAG 490
465 GGGATGCCAATGATCCCGAGACTTTTGGACAGCTGGGATCTCCACAGAGAGAG 524
491 GGGATGCCAATGATCCCGAGACTTTTGGACAGCTGGGATCTCCACAGAGAGAG 550
525 GGGATGCCAATGATCCCGAGACTTTTGGACAGCTGGGATCTCCACAGAGAGAG 584
551 TGGATTCAGATTTATTCAGAGCAAGGAGTTCACCACTGCAAGAGTATTCACCATCA 610
585 TGGATTCAGATTTATTCAGAGCAAGGAGTTCACCACTGCAAGAGTATTCACCATCA 644
611 AAACCAACCTGACAGACAGAGATATATATGCTGGGATCCAGAGATCTTCTC 670
645 AAACCAACCTGACAGACAGAGATATATATGCTGGGATCCAGAGATCTTCTC 704
671 TGAATGCGCTGCTCCATGCTGATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 730
705 TGAATGCGCTGCTCCATGCTGATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 764
731 CCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 790
765 CCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 824
791 GCTACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 850
825 GCTACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 884
851 TACACCTTTGGAGATGACATATTTACTGATGATGATGATGATGATGATGATGATGAT 910
885 TACACCTTTGGAGATGACATATTTACTGATGATGATGATGATGATGATGATGATGAT 944
911 TTAGGAGAGAACCCAGACTTTTCCCAAGCTGAGATTTTGGAAAGTGGAAATG 970
945 TTAGGAGAGAACCCAGACTTTTCCCAAGCTGAGATTTTGGAAAGTGGAAATG 1004
971 ATGACTTCAAGCTGAGATCTTCAAGTGAAGGATGATGATGATGATGATGATGATGAT 1030
1005 ATGACTTCAAGCTGAGATCTTCAAGTGAAGGATGATGATGATGATGATGATGATGAT 1064
1031 TGAATATGCTGTTTGGGCTTTCAAGAGAGCTTGAAGATGATGATGATGATGATGATGAT 1090
1065 TGAATATGCTGTTTGGGCTTTCAAGAGAGCTTGAAGATGATGATGATGATGATGATGAT 1124
1091 GTTGGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1150
1125 GTTGGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1184
1151 AAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1210
1185 AAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1244
1211 ACTGTCAGATCTTCCATTAATAAGGCTTTGATTAATCTCACTGAGAGGATCTGACA 1270
1245 ACTGTCAGATCTTCCATTAATAAGGCTTTGATTAATCTCACTGAGAGGATCTGACA 1304
1271 ATGCTGAGTTTGAACAAAGTGAAGAGATGAAATGATGATGATGATGATGATGATGAT 1330
1305 ATGCTGAGTTTGAACAAAGTGAAGAGATGAAATGATGATGATGATGATGATGATGAT 1364
1331 TATGCTGATTTCAATCCACGATCTTATTAAGAGGATGATGATGATGATGATGATGATGAT 1390
1365 TATGCTGATTTCAATCCACGATCTTATTAAGAGGATGATGATGATGATGATGATGATGAT 1424

1391 TGAATATTTTGAATATTTTGAATTTTCAAGGCTATCCCTCAATCTGAGGAGC 1450
1425 TGAATATTTTGAATATTTTGAATTTTCAAGGCTATCCCTCAATCTGAGGAGC 1484
1451 TGAATATTTTGAATATTTTGAATTTTCAAGGCTATCCCTCAATCTGAGGAGC 1504
1485 TGAATATTTTGAATATTTTGAATTTTCAAGGCTATCCCTCAATCTGAGGAGC 1544
1505 TGAATATTTTGAATATTTTGAATTTTCAAGGCTATCCCTCAATCTGAGGAGC 1536
1545 TGAATATTTTGAATATTTTGAATTTTCAAGGCTATCCCTCAATCTGAGGAGC 1576

RESULT 15

US-09-652-121-6686
Sequence 6686, Application US/09652121
GENERAL INFORMATION:
APPLICANT: Diselano, Peter
TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
FILE REFERENCE: 1600.1188-001
CURRENT APPLICATION NUMBER: US/09/652,121
CURRENT FILING DATE: 2000-08-30
PRIOR APPLICATION NUMBER: 60/151,129
PRIOR FILING DATE: 1999-08-30
NUMBER OF SEQ ID NOS: 7615
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6686
LENGTH: 1829
TYPE: DNA
ORGANISM: Homo sapiens
US-09-652-121-6686

Query Match 98.0%; Score 1505.4; DB 25; Length 1829;
Best Local Similarity 99.3%; Pred. No. 0;

Matches 1522; Conservative 0; Mismatches 4; Indels 6; Gaps 1;

11 GAGCACTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 70
45 GAGCACTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 104
71 CACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 130
105 CACTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 164
131 GCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 190
165 GCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 224
191 GCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 250
225 GCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 284
251 ACCGCAAGGCAACCGGCAACCGGCAACCGGCAACCGGCAACCGGCAACCGGCAACCGGCA 310
285 ACCGCAAGGCAACCGGCAACCGGCAACCGGCAACCGGCAACCGGCAACCGGCAACCGGCA 344
311 ATGAAATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 370
345 ATGAAATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 404
371 TGTGTTTATCAAGGATTCACAAATGCTAAAGAGCTCTTCCAGAGGATGAAATCT 430
405 TGTGTTTATCAAGGATTCACAAATGCTAAAGAGCTCTTCCAGAGGATGAAATCT 464
431 GGGATGCCAATGATCCCGAGACTTTTGGACAGCTGGGATCTCCACAGAGAGAG 490
465 GGGATGCCAATGATCCCGAGACTTTTGGACAGCTGGGATCTCCACAGAGAGAG 524
491 GGGATGCCAATGATCCCGAGACTTTTGGACAGCTGGGATCTCCACAGAGAGAG 550
525 GGGATGCCAATGATCCCGAGACTTTTGGACAGCTGGGATCTCCACAGAGAGAG 584

QY 551 TGGATCAGATTATTTCAGGACAGGGAGTTACCAATGGAAGAAGTATGACCATCA 610
| | | | |
Db 585 TGAATCAGATTATTTCAGGACAGGGAGTTACCAATGGAAGAAGTATGACCATCA 644
| | | | |
QY 611 AATCCAACTCCGACGACGACGAGAAATCATCATGTCGCTTGGAAATCAGAGATCTTCCCTC 670
| | | | |
Db 645 AATCCAACTCCGACGACGAGAAATCATCATGTCGCTTGGAAATCAGAGATCTTCCCTC 704
| | | | |
QY 671 TGAATGAGCTGCTCCATGACATGCTCCCTGACAGTTCTATGTGTGAACATGAGCTGT 730
| | | | |
Db 705 TGAATGAGCTGCTCCATGACATGCTCCCTGACAGTTCTATGTGTGAACATGAGCTGT 764
| | | | |
QY 731 CCTGCCAGCTGTACAGAGATGAGGAGACATGAGGCTCTGCTGTCCTTTCAATGCGCA 790
| | | | |
Db 765 CCGCCAGCTGTACAGAGATGAGGAGACATGAGGCTCTGCTGTCCTTTCAATGCGCA 824
| | | | |
QY 791 GCTACGCTCTGCTCACTGACATGATTCGCGACATCAAGGCTCTGAGCCAGGTCCTTA 850
| | | | |
Db 825 GCTACGCTCTGCTCACTGACATGATTCGCGACATCAAGGCTCTGAGCCAGGTCCTTA 884
| | | | |
QY 851 TACACCTTTGGGAGATGACATTTTACCTGATCATGACATGACATGACATGACATGAC 910
| | | | |
Db 885 TACACCTTTGGGAGATGACATTTTACCTGATCATGACATGACATGACATGACATGAC 944
| | | | |
QY 911 TTTACGAGAGACCCGACCTTTCCCAAGCTCAGATTTCTGAAAAGTTGAGAAAATTG 970
| | | | |
Db 945 TTTACGAGAGACCCGACCTTTCCCAAGCTCAGATTTCTGAAAAGTTGAGAAAATTG 1004
| | | | |
QY 971 ATGACTTCAAAGCTGAGACCTTTCAAGTTGAGAGGTACATCCGATCCAACTATTAAA 1030
| | | | |
Db 1005 ATGACTTCAAAGCTGAGACCTTTCAAGTTGAGAGGTACATCCGATCCAACTATTAAA 1064
| | | | |
QY 1031 TGGAAATGCTGTTTAAAGGCTCTTCAAGAGCTGGAAGATATGTGACGCTTTAAGG 1090
| | | | |
Db 1065 TGGAAATGCTGTTTAAAGGCTCTTCAAGAGCTGGAAGATATGTGACGCTTTAAGG 1124
| | | | |
QY 1091 GTTGGGCTGGATGCCAGGTAAAGTCTTTTGTCTTAAAGAAAGAACTAGAGTCA 1150
| | | | |
Db 1125 GTTGGGCTGGATGCCAGGTAAAGTCTTTTGTCTTAAAGAAAGAACTAGAGTCA 1184
| | | | |
QY 1151 AAAATCTGTCCGTGACCTATCATATAATAAAGGCTTTGAGTAACTCACTGAGGTAATGTA 1210
| | | | |
Db 1185 AAAATCTGTCCGTGACCTATCATATAATAAAGGCTTTGAGTAACTCACTGAGGTAATGTA 1244
| | | | |
QY 1211 ACTGTCCAGTCTTTTCCATATAATAAAGGCTTTGAGTAACTCACTGAGGTAATGTA 1270
| | | | |
Db 1245 ACTGTCCAGTCTTTTCCATATAATAAAGGCTTTGAGTAACTCACTGAGGTAATGTA 1304
| | | | |
QY 1271 ATGCTGAGGTATGAACAAAGTGAAGAAATGAATATGCTCTTACCAAAAACATG 1330
| | | | |
Db 1305 ATGCTGAGGTATGAACAAAGTGAAGAAATGAATATGCTCTTACCAAAAACATG 1364
| | | | |
QY 1331 TATGTGATTTCAATCCACGACTTATAAAGAGGTGAGTAATTCACAGCTATTTT 1390
| | | | |
Db 1365 TATGTGATTTCAATCCACGACTTATAAAGAGGTGAGTAATTCACAGCTATTTT 1424
| | | | |
QY 1391 TGAATATTTTAAATATTTTAAAGATTTCAAACTATTTCCCTCAATCTGAGGAGC 1450
| | | | |
Db 1425 TGAATATTTTAAATATTTTAAAGATTTCAAACTATTTCCCTCAATCTGAGGAGC 1484
| | | | |
QY 1451 TGAATATTTTAAATATTTTAAAGATTTCAAACTATTTCCCTCAATCTGAGGAGC 1504
| | | | |
Db 1485 TGAATATTTTAAATATTTTAAAGATTTCAAACTATTTCCCTCAATCTGAGGAGC 1544
| | | | |
QY 1505 TTAATGTTGCTATTAATAAAGAGTGTCTGC 1536
| | | | |
Db 1545 TTAATGTTGCTATTAATAAAGAGTGTCTGC 1576
| | | | |

QY 927 ACCCTTTCCTCAAGCTCAG

```

DB 15630 ACCCTTCCCAAAGCTCAGGATCTTCGAAAAGTTGAGAAAATGATGACTTCAAGCTGA 15689
QY 987 AGACTTTCAGATTGAAGGGTACATCCGCATCCACTATTAAATGGAATGGCTGTTTA 1046
DB 15690 AGACTTTCAGATTGAAGGGTACATCCGCATCCACTATTAAATGGAATGGCTGTTTA 15749
QY 1047 GGGTGCCTTTCAGAAAGCTNAGAAAGATTTGTCTGCTTTAGGGGTTGGGCTGATGCCG 1106
DB 15750 GGGTGCCTTTCAGAAAGCTCAGAGGATTTGTCTGCTTTAGGGGTTGGGCTGATGCCG 15809
QY 1107 AGGTAAAGTCTTTTGGCTCTAAAGAAAAGAACTAGGTCAAAAATCTGCTCCGTAC 1166
DB 15810 AGGTAAAGTCTTTTGGCTCTAAAGAAAAGAACTAGGTCAAAAATCTGCTCCGTAC 15869
QY 1167 CTATCAGTTTATTTTATTTTAAAGATTTGCACTGCAATGTAATGTCAGTTCTTT 1226
DB 15870 CTATCAGTTTATTTTATTTTAAAGATTTGCACTGCAATGTAATGTCAGTTCTTT 15929
QY 1227 CCATATATAAGGCTTTGAGTTAATCACTGAGGGTATCTGACAAATGCTGAGTTTAA 1286
DB 15930 CCATATATAAGGCTTTGAGTTAATCACTGAGGGTATCTGACAAATGCTGAGTTTAA 15989
QY 1287 CAAAGTGAAGAAATGAATGTATGTCTTTAGGAAAAATGTAATGTGCAATTCATC 1346
DB 15990 CAAAGTGAAGAAATGAATGTATGTCTTTAGGAAAAATGTAATGTGCAATTCATC 16049
QY 1347 CCAGTACTTATAAGAGAGTTGTGGAATTCACACACTATTTTGGAAATATTTTAA 1406
DB 16050 CCAGTACTTATAAGAGAGTTGTGGAATTCACACACTATTTTGGAAATATTTTAA 16109
QY 1407 TATTTTAAAGATTTTCAAGCTATTCCTCAATCTGAGGAGCTGAGTAAACCATGGA 1466
DB 16110 TATTTTAAAGATTTTCAAGCTATTCCTCAATCTGAGGAGCTGAGTAAACCATGGA 16169
QY 1467 TCATGATGTAGAGTGTGTTATGAACT-----TTAAGTTGTTTATATGTTGCTATA 1520
DB 16170 TCATGATGTAGAGTGTGTTATGAACTTTAAGTTATATGTTTATATGTTGCTATA 16229
QY 1521 TAAAGAGTGTCTGC 1536
DB 16230 TAAAGAGTGTCTGC 16245

```

RESULT 2
US-09-954-531-348 Application US/09954531

Sequence 348 Patent No. US20020165180A1

GENERAL INFORMATION:

APPLICANT: Weaver, Zoe

TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cans

TITLE OF INVENTION: Gene Sets

FILE REFERENCE: 689290-77

CURRENT APPLICATION NUMBER: US/09/954,531

CURRENT FILING DATE: 2002-05-02

PRIOR APPLICATION NUMBER: US/60/233,133

PRIOR FILING DATE: 2000-09-18

PRIOR APPLICATION NUMBER: US/60/234,009

PRIOR FILING DATE: 2000-09-20

PRIOR APPLICATION NUMBER: US/60/234,034

PRIOR FILING DATE: 2000-09-20

PRIOR APPLICATION NUMBER: US/60/234,509

PRIOR FILING DATE: 2000-09-22

PRIOR APPLICATION NUMBER: US/60/234,567

PRIOR FILING DATE: 2000-09-22

NUMBER OF SEQ ID NOS: 1392

SOFTWARE: PatentIn version 3.0

SEQ ID NO 348

LENGTH: 18596

TYPE: DNA

ORGANISM: Homo sapiens

US-09-954-531-348

```

Query Match 39.8%; Score 611.8; DB 9; Length 18596;
Best Local Similarity 90.4%; Pred. No. 2.5e-154;
Matches 665; Conservative 0; Mismatches 65; Indels 6; Gaps 1;

QY 807 GTACATGATTTGGCACAATCAGAGGGCTGAGCCAGGTGACTTTATATACACTTTGGAGGA 866
DB 15510 GAACTTTGTTGATCAGATCTGTGTACTTTGTTTCAGGACATGAGAGACATTTACACAG 15569
QY 867 TGCAATATTTTACCTGAATCAGATCAGCCAGTAAATTTAGCTTACGAGGAGACCCAG 926
DB 15570 GTCTGATCAATTTAGGAAATATATAGCTTATTTTGTTTTACCTTACGAGGAGACCCAG 15629
QY 927 ACCTTTCCAAAGCTCAGATTTCTGAAAAGTTGAGAAAATTTGATGACTTCAAACTGA 986
DB 15630 ACCTTTCCAAAGCTCAGATTTCTGAAAAGTTGAGAAAATTTGATGACTTCAAACTGA 15689
QY 987 AGACTTTCAGATTGAAGGGTACATCCGCATCCACTATTAAATGGAATGGCTGTTTA 1046
DB 15690 AGACTTTCAGATTGAAGGGTACATCCGCATCCACTATTAAATGGAATGGCTGTTTA 15749
QY 1047 GGGTGCCTTTCAGAAAGCTNAGAAAGATTTGTCTGCTTTAGGGGTTGGGCTGATGCCG 1106
DB 15750 GGGTGCCTTTCAGAAAGCTCAGAGGATTTGTCTGCTTTAGGGGTTGGGCTGATGCCG 15809
QY 1107 AGGTAAAGTCTTTTGGCTCTAAAGAAAAGAACTAGGTCAAAAATCTGCTCCGTAC 1166
DB 15810 AGGTAAAGTCTTTTGGCTCTAAAGAAAAGAACTAGGTCAAAAATCTGCTCCGTAC 15869
QY 1167 CTATCAGTTTATTTTATTTTAAAGATTTGCACTGCAATGTAATGTCAGTTCTTT 1226
DB 15870 CTATCAGTTTATTTTATTTTAAAGATTTGCACTGCAATGTAATGTCAGTTCTTT 15929
QY 1227 CCATATATAAGGCTTTGAGTTAATCACTGAGGGTATCTGACAAATGCTGAGTTTAA 1286
DB 15930 CCATATATAAGGCTTTGAGTTAATCACTGAGGGTATCTGACAAATGCTGAGTTTAA 15989
QY 1287 CAAAGTGAAGAAATGAATGTATGTCTTTAGGAAAAATGTAATGTGCAATTCATC 1346
DB 15990 CAAAGTGAAGAAATGAATGTATGTCTTTAGGAAAAATGTAATGTGCAATTCATC 16049
QY 1347 CCAGTACTTATAAGAGAGTTGTGGAATTCACACACTATTTTGGAAATATTTTAA 1406
DB 16050 CCAGTACTTATAAGAGAGTTGTGGAATTCACACACTATTTTGGAAATATTTTAA 16109
QY 1407 TATTTTAAAGATTTTCAAGCTATTCCTCAATCTGAGGAGCTGAGTAAACCATGGA 1466
DB 16110 TATTTTAAAGATTTTCAAGCTATTCCTCAATCTGAGGAGCTGAGTAAACCATGGA 16169
QY 1467 TCATGATGTAGAGTGTGTTATGAACT-----TTAAGTTGTTTATATGTTGCTATA 1520
DB 16170 TCATGATGTAGAGTGTGTTATGAACTTTAAGTTATATGTTTATATGTTGCTATA 16229
QY 1521 TAAAGAGTGTCTGC 1536
DB 16230 TAAAGAGTGTCTGC 16245

```

RESULT 3
US-09-880-107-1590

Sequence 1590 Patent No. US20020142981A1

GENERAL INFORMATION:

APPLICANT: Horne, Darci T.

APPLICANT: Vockley, Joseph G.

APPLICANT: Scherf, Uwe

APPLICANT: Gene Logic, Inc.

TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer

FILE REFERENCE: 44921-5028-WO

CURRENT APPLICATION NUMBER: US/09/880,107

CURRENT FILING DATE: 2001-06-14

PRIOR APPLICATION NUMBER: US 60/211,379

PRIOR FILING DATE: 2000-06-14

PRIOR APPLICATION NUMBER: US 60/237,054

Qy 1467 TCATGATGCAAGTGGGTATGAAC-----TAAAGTGGTTTATGCTGTATAA 1520
Db 16170 TCATGATGCAAGTGGGTATGAACTTAAAGTTAAGTTGTTTATAGTTGCTATAA 16222
Qy 1521 TAAAGAGTGTTCGC 1536

Db 16230 TAAAGAGGTCTTCGC 16245

RESULT 5

US-10-046-935-2148
; Sequence 2148, Application US/10046935
; Patent No. US20020156011A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugui
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Secrist, Heather
; APPLICANT: Wang, Aljun
; APPLICANT: Stolck, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.527C1
; CURRENT APPLICATION NUMBER: US/10/046,935
; CURRENT FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 2239
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2148
; LENGTH: 532
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-046-935-2148

Query Match 34.6%; Score 532; DB 9; Length 532;
Best Local Similarity 100.0%; Pred. No. 9.3e-134;
Matches 532; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 203 ACCTGGGGGAGATCCCAACATCCCGCGCGGTGAGGAGGACGACGCGGCA 262
Db 1 ACCTGGGGGAGATCCCAACATCCCGCGCGGTGAGGAGGACGACGCGGCA 60
QY 263 CCGGACCCCTGTGGATTCGGGATGCGAGGCGCTACAGTGAAGATGATTCCTC 322
Db 61 CCGGACCCCTGTGGATTCGGGATGCGAGGCGCTACAGTGAAGATGATTCCTC 120
QY 323 TGCTGACAAACCAACGTTGTTCTGAAAGGGTGTGGAAGAGTTGCTGTGTTATCA 382
Db 121 TGCTGACAAACCAACGTTGTTCTGAAAGGGTGTGGAAGAGTTGCTGTGTTATCA 180
QY 383 AGGGATCCCAAAATGCTAAAGAGCTGTCTTCCAGGAGTGAATAATCTGGATGCAATG 442
Db 181 AGGGATCCCAAAATGCTAAAGAGCTGTCTTCCAGGAGTGAATAATCTGGATGCAATG 240
QY 443 GATCCCGAGACTTTTGGACAGCTGGGATTTCTCACAGAGAAAGGGGACTTGGGCC 502
Db 241 GATCCCGAGACTTTTGGACAGCTGGGATTTCTCACAGAGAAAGGGGACTTGGGCC 300
QY 503 CAGTTTATGGCTTCCAGTGGAGGCAATTTGGGGGCAAGATCAGAGATGATCAGATT 562
Db 301 CAGTTTATGGCTTCCAGTGGAGGCAATTTGGGGGCAAGATCAGAGATGATCAGATT 360
QY 563 ATTGAGGACAGGAGTGAACCAATGCAAGAGATGTTGACACATCAAAACCAACCTTG 622
Db 361 ATTGAGGACAGGAGTGAACCAATGCAAGAGATGTTGACACATCAAAACCAACCTTG 420
QY 623 AGGACAGAGATCATCATGCGCTTGGATCCAGAGATCTTCTCTGATGGCGCTGC 682
Db 421 AGGACAGAGATCATCATGCGCTTGGATCCAGAGATCTTCTCTGATGGCGCTGC 480
QY 683 CTCATGCGATGCCCTCTCTGCGCAATTTATGTGTGAACAGTAGCTGTCTTG 734
Db 481 CTCATGCGATGCCCTCTCTGCGCAATTTATGTGTGAACAGTAGCTGTCTTG 532

RESULT 6
US-09-922-217-443/C
; Sequence 443, Application US/09922217
; Patent No. US20020076414A1
; GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stolck, John A.
; APPLICANT: Wang, Tonglong
; APPLICANT: Jiang, Yugui
; APPLICANT: Smith, Carole Lynn
; APPLICANT: King, Gordon B.
; APPLICANT: Wang, Aljun
; APPLICANT: Clapper, Jonathan D.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; TITLE OF INVENTION: OF COLON CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.471C13
; CURRENT APPLICATION NUMBER: US/09/922,217
; CURRENT FILING DATE: 2001-08-03
; NUMBER OF SEQ ID NOS: 1124
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 443
; LENGTH: 346
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-922-217-443

Query Match 22.2%; Score 340.8; DB 10; Length 346;
Best Local Similarity 98.8%; Pred. No. 2.4e-82;
Matches 342; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1007 ACATCCGATCCACTATTAATGGAATGCTGTTAGGAGTGTTCAGAGAGCTN 1066
Db 346 ACATCCGATCCACTATTAATGGAATGCTGTTAGGAGTGTTCAGAGAGCTN 287
QY 1067 GAAGATATGTCAGTCTTAGGGCTGGCTGATCCGAGGTAAAGTCTTTTCT 1126
Db 286 GAAGATATGTCAGTCTTAGGGCTGGCTGATCCGAGGTAAAGTCTTTTCT 227
QY 1127 CTAAAGAAAGAGACTAGTCAAAAATCTGTCGTAATCAAGTATTAATTTTA 1186
Db 226 CTAAAGAAAGAGACTAGTCAAAAATCTGTCGTAATCAAGTATTAATTTTA 167
QY 1187 AGGATGTCACACTGGCAATGTAATCTGTGTCAGTCTTTCATTAATTAAGCTTTGAG 1246
Db 166 AGGATGTCACACTGGCAATGTAATCTGTGTCAGTCTTTCATTAATTAAGCTTTGAG 107
QY 1247 TTACTCATGAGGTATCTGACATGCTGAGCTTATGAAAGAGAGATGAAT 1306
Db 106 TTACTCATGAGGTATCTGACATGCTGAGCTTATGAAAGAGAGATGAAT 47
QY 1307 GTATGTGCTTTAGCAAAAACATGATGTGCAATTCATCCACGT 1352
Db 46 GTATGTGCTTTAGCAAAAACATGATGTGCAATTCATCCACGT 1

RESULT 7
US-09-833-263-443/C
; Sequence 443, Application US/09833263
; Patent No. US20020110547A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Aljun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Stolck, John A.
; APPLICANT: Meagher, Madeleine J.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.471C12
; CURRENT APPLICATION NUMBER: US/09/833,263
; CURRENT FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 1093
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 443
; LENGTH: 346
; TYPE: DNA

ORGANISM: Homo sapien
US-09-833-263-443
Query Match 22.2%; Score 340.8; DB 10; Length 346;
Best Local Similarity 98.8%; Pred. No. 2.4e-82;
Matches 342; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1007 ACATCCGATCCACTATTAAATGAAATGCTGTTTAAAGGCTCTTCAAAAGGCTTN 1066
DB ACAATCCGATCCACTATTAAATGAAATGCTGTTTAAAGGCTCTTCAAAAGGCTTT 287
QY 1067 GAAGGATATGTGAGCTTTAGGGGTGGGCTGGATGCCAGGTAAGTCTTTTGT 1126
DB GAAGGATATGTGAGCTTTAGGGGTGGGCTGGATGCCAGGTAAGTCTTTTGT 227
QY 1127 CTAAAGANAAGAACTAGTCAAAATCTGTCCTGACCTATGATTAATTTTAA 1186
DB CTAAAGANAAGAACTAGTCAAAATCTGTCCTGACCTATGATTAATTTTAA 167
QY 1187 AGATGTTGCCACTGSCAAATGTAACTGTGCCAGTCTTTCATTAATAAAGCTTTGAG 1246
DB AGATGTTGCCACTGSCAAATGTAACTGTGCCAGTCTTTCATTAATAAAGCTTTGAG 107
QY 1247 TTAATCACTGAAGGATCTGAATGCTGAAGTTTAAGAAAGTGAAGATGAAT 1306
DB TTAATCACTGAAGGATCTGAATGCTGAAGTTTAAGAAAGTGAAGATGAAT 47
QY 1307 GTATGTCTCTTAGCAAAACATGTATGCTATTCAATCCACGT 1352
DB GTATGTCTCTTAGCAAAACATGTATGCTATTCAATCCACGT 1

RESULT 8
US-09-741-669-181
Sequence 181, Application US/09741669
Patent No. US2002022718A1
GENERAL INFORMATION:
APPLICANT: Forsyth, R. Allen
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
TITLE OF INVENTION: Genes identified as required for
FILE REFERENCE: ELITRA.009A
CURRENT APPLICATION NUMBER: US/09/741,669
PRIOR FILING DATE: 2000-12-19
PRIOR APPLICATION NUMBER: US 60/173005
PRIOR FILING DATE: 1999-12-23
NUMBER OF SEQ ID NOS: 481
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 181
LENGTH: 795
TYPE: DNA
ORGANISM: Escherichia coli
REMARKS:
NAME/KEY: CDS
LOCATION: (1)...(795)
US-09-741-669-181

Query Match 11.7%; Score 180; DB 10; Length 795;
Best Local Similarity 54.0%; Pred. No. 8.6e-39;
Matches 458; Conservative 0; Mismatches 330; Indels 60; Gaps 2;

QY 199 CAGTACCTGGGCGAGATCCACACATCTCCGCTGGCGGTCAAGAGAGACGACGCGC 258
DB CAGTATTTAGACTGATGCAAAAGTCTGACAGAGGACAGCAAAAGACGCGTACC 66
QY 259 GGCACCGGACCTGTGCTGATTCGCGATCGACAGCGCGGTACAGGCTGAGATGAATTC 318
DB GGAACCGGACGCTTTCATTTTGTGCATCAGATCCGTTTAACTGCAAGATGATTC 126
QY 319 CCTGCTGACACACAAAGTGTGTCTGGAAGGGTGTGTTGAGAGAGTGTGCTGCTTT 378
DB CGCTGTGTCACATTAAGCTTTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTT 186

QY 379 ATCAAGGATCCACAAATGCTAAAGAGCTGTCTTCAAGGAGTGAATTTGAGATGCC 438
DB CTGCGAGGCGACACTTAACATCTGATCTACACAAACATGCTCACTGAGGACGA 246
QY 439 AATGATCCGACACTTTTGAACAGCTGGGATTTCTCCACCAAGAAAGGAGCTTG 498
DB TGGG-----CCGATGAAACCGGACCTC 270
QY 499 GGCCAGTTTATGAGCTTCCAGTGAAGCATTTTGGGCGACAAATACAGATATGAAATCA 558
DB GGGCCAGTGTATGTTAAAGAGTGGCGCTGGCAACGCGAGAT----- 315
QY 559 GATTAATCAGACAGGAGTTGACCACTGCAAGATGATTAAGACCATCAAAACCAAC 618
DB -----GGTCTCAATTAAGACGATCACTAGCTGTAAGACAGCTGAAAGAGAC 366
QY 619 CCTGACGACAAATATCATATGCTGCTGATCATCAAGATTTCTGTATGCGC 678
DB CCGATTTGCGCGCATTAATGTTTCACTGGAACGTAGCGCACTGATTAATATGCGC 426
QY 679 CTGCTTCATSCATGCGCTCTGCCAGTTCTATGCTGAACAGTGAAGCTGCTGCGCAG 728
DB CTGCGACCTGCGCATGATTTTCCAGTTCTATGCTGAACAGGCAACTCTTTCAG 486
QY 739 CTGACAGAGATGCGAGACATGCGCTGCTGCTGCTTTCATCATGCGCAGTACGCC 798
DB CTGTAATGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 546
QY 799 CTGCTGATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 858
DB TATTTGATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 606
QY 859 TTGGAGATGACATTAATTAATGATCATGATGATGATGATGATGATGATGATGAT 918
DB GGTGCGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 666
QY 919 GAACCGACCTTTCCAAAGCTCAGATTTCTGCAAAAGTTGAGAAATGATGATCTTC 978
DB GAACCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 726
QY 979 AAGGTGAAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1038
DB GCTTTGCAAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 786
QY 1039 GCTGTTA 1046
DB GCTATCTA 794

RESULT 9
US-09-790-988-1
Sequence 1, Application US/09790988
Patent No. US20020127687A1
GENERAL INFORMATION:
APPLICANT: SHIGEMOTO, SHUJI
APPLICANT: WATANABE, HIDEMI
APPLICANT: HATORI, MASAHIRA
APPLICANT: SAKAKI, YOSHIYUKI
TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
FILE REFERENCE: 081356/0159
CURRENT APPLICATION NUMBER: US/09/790,988
PRIOR FILING DATE: 2001-02-23
PRIOR APPLICATION NUMBER: JP2000-107160
PRIOR FILING DATE: 2000-04-07
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 640681
TYPE: DNA
ORGANISM: Buchnera sp.
US-09-790-988-1

Query Match 9.0%; Score 137.6; DB 10; Length 640681;
 Best Local Similarity 55.4%; Pred. No. 1.3e-25;
 Matches 266; Conservative 0; Mismatches 214; Indels 0; Gaps 0;

QY 567 AGGACAGGAGTTGACCACTGCAAGAGTGTGACCAATCAAAAACCAACCTGACGA 626
 Db 480465 AGGACAGCAATGATCAATAATAATGATTTAATCAATTAATAAAAAATCCGATTC 480524
 QY 627 CAGAAAGATCATCATGTGCGCTTGGATTCAGAGATCTTCTCTGATGGCGCTGCTCC 666
 Db 480525 CCGTAGATGTAGTTTCTAGTTGAGATGTTGAGATATGATTAATAAGATTACCTCC 480584
 QY 687 ATGCGATGCCCCCTGCGCACTTCTATGTGTGAACAAGAGCTGTCTCGACGCTGTACA 746
 Db 480585 TTGCGATGTCTATTTCAATTTATATGATTAACAATACATTAAGTTGCAATGTACCA 480644
 QY 747 GAGATCGGAGACATGGGCGCTCGGTGTGCTTTCAACATGCCAGCTAGCGCTGCTCAC 806
 Db 480645 GCGTTCTTGATGTGTTTCTAGGACTACCTTTTAATAGCTAGTACTCAATACTTAT 480704
 QY 807 GATCATGATTTGGACATCAAGGCGCTGAAGCCAGTACTTATACAACATTGAGGAGA 866
 Db 480705 AACATGATAGCAACAACAAATGATTTAAAGTTGAGATTTTGTGACAGAGGCGA 480764
 QY 867 TGCAATATTTACCTGATCATCATGAGCGCACTGAAAATTCAGCTTCAGCGAGAACCCAG 926
 Db 480765 TGTTCATCTTAT 480824
 QY 927 AACTTTCCCAAGCTCAGATTTCTTGAAAAGTTGAGAAAATGATGACTTCAAGCTGA 986
 Db 480825 AACGTTCCCAACTAACATCTCTCAAAAACCTCAGTCAATATTTCAATATGTTTGA 480884
 QY 987 AGACTTTCAGTTGAGAGGTTACATTCGCAATCCATATTAATAAGAAATGGCTGTTA 1046
 Db 480885 AGACTTTAAATATATGATATCATCTTATCTGCTATCAAGAGGAAATATCATATA 480944

RESULT 10

US-09-954-197-1
 ; Sequence 1, Application US/09954197
 ; Patent No. US20020107379A1
 ; GENERAL INFORMATION:
 ; APPLICANT: MARK, Achim
 ; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES CODING FOR THE thyA GENE
 ; FILE REFERENCE: 032301 WD 201
 ; CURRENT APPLICATION NUMBER: US/09/954,197
 ; CURRENT FILING DATE: 2001-12-18
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 1
 ; LENGTH: 1200
 ; TYPE: DNA
 ; ORGANISM: Corynebacterium glutamicum
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (231)..(1028)
 ; OTHER INFORMATION:
 ; US-09-954-197-1

Query Match 8.3%; Score 126.8; DB 10; Length 1200;
 Best Local Similarity 56.7%; Pred. No. 2.4e-24;
 Matches 254; Conservative 0; Mismatches 191; Indels 3; Gaps 1;

QY 578 TTGACCAACTGCAAAAGATGATTTGACCATCAAAACCACTGACGACAGAAATCA 637
 Db 565 TTGACCAAGATCTCAGGTGCTTTAGAAACTCTGCGAAACACCTGATTCACCTGCAATA 624
 QY 638 TCAATGCGCTTGGAATCCAAGAGATCTTCCCTGATGCGCGCTCCATGCCATGCGCC 697
 Db 625 TTGCTCGCGGTGGAATGTTTCCGAGCTGTAAGAAACATGAGCTTTCCCTTGTCACTTGC 684
 QY 698 TCTGCAAGTCTATGTTGTGAACAAGTGAAGCTGTCTCTGCAAGCTGTATCAAGATCGGAG 757

Db 685 TTTTCAGCTCTATGTGCGCATGAGCAACTGTTTGCAGCTTACCAAGCTTCTGCGG 744
 QY 758 ACATGGGCGCTGGGTGCGCTTTCAACATCGCAGTACGACCTGCTACATATGATTT 817
 Db 745 ACATGTTCTGGGTGCGCTTTCACATCCATCTTATATCACTGCTTCCACCAATGTTTG 804
 QY 818 CGCATCAAGGCGCTGGAAGCGAGTGAATTTATACACTTTTGGAGATGACATATTT 877
 Db 805 CCGACGAGAGGCTTGAAGTGGGCGAGATTTGAGATCTGGGGGAGATGCGCAATTT 864
 QY 878 ACTGAATCATCGACCACTGAAAATTCAGCTTACGAGAGAACCCAGACTTTCCCA 937
 Db 865 ATGACAAACCAAGAAAGAGTGGCGAGCAGCTGAGCCGCGAACTGCGCCCTTACCCCA 924
 QY 938 AGCTCAGATTTCTGAAAAGTTGAGAAATATATATGACTTCAAACTGTAAGACTTTGAG 997
 Db 925 ---CCTTGAGCTCAACAAGGCGAGCTCCATGTTTGAATGAGCTTGATGATCAATCACG 981
 QY 998 TTGAAGGTACATTCGCACTCCACTAT 1025
 Db 982 TGTGCGGTACGATCCACACCATTTGAT 1009

RESULT 11

US-09-987-446-1
 ; Sequence 1, Application US/09987446
 ; Patent No. US20020115162A1
 ; GENERAL INFORMATION:
 ; APPLICANT: FARWICK, Mike, et al.
 ; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES CODING FOR THE cysQ GENE
 ; FILE REFERENCE: 032301 WD 245
 ; CURRENT APPLICATION NUMBER: US/09/987,446
 ; CURRENT FILING DATE: 2001-11-14
 ; NUMBER OF SEQ ID NOS: 2
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 1
 ; LENGTH: 2730
 ; TYPE: DNA
 ; ORGANISM: Corynebacterium glutamicum
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1014)..(1769)
 ; OTHER INFORMATION:
 ; US-09-987-446-1

Query Match 8.3%; Score 126.8; DB 10; Length 2730;
 Best Local Similarity 52.0%; Pred. No. 3.9e-24;
 Matches 407; Conservative 0; Mismatches 312; Indels 63; Gaps 3;

QY 244 AAGGACACCGGACGAGCGACCGGACCGCTGCGATTTGGCATGAGCGCGCTACAGC 303
 Db 1851 AAGGACGACCGGACCGGACCGGACCGCTGCTTATTTGGAACAATAATCCGCTTTGAT 1910
 QY 304 CTGAGAGATGATTCCTCTGCTGACCAACCAAGTGTCTTGAAGAGGTGTTTGAG 363
 Db 1911 CTGATGAAAGTTTCCCTTCTGACCAAGAAAGTTCATTTCACTCTGTGTGGGT 1970
 QY 364 GAGTGTCTGTTTATCAGAGATTCACAAATGCTTAAAGCTGTCTTCCAAAGGAGTG 423
 Db 1971 GAGCTTTGTGTCTTCCCTTCAAGGGGATTCACATGCTCAAAATGCGAGATTAACAATC 2030
 QY 424 AAAATCTGGAGTGCATATGATCCCGAGATCTTTTGAACAGCTTGGATTTCCACAGA 483
 Db 2031 CGATTTGGAATGAATGGCAGAT----- 2054
 QY 484 GAAAGAGGAGCTTGGGCGCAATTTATGCTTCCAGTGAAGAGATTTTGGGCGAGATAC 543
 Db 2055 GAGGACGCGGAGCTGGGCGCTGTTATGTGTCCAGTGGGCTTCTTGGC----- 2103
 QY 544 AAGATATGAAATCAATTTATTCAGGACAGGAGTGAACCACTGCAAAAGATGATGAC 603
 Db 2104 -----CAACCTGATGCTGTCACTATGACAGATCTCAGGTGCTTTAGAA 2150

QY	604	ACCAAAACCAACCC	TGACGACAGAAAGATTCATGTGCGTGTGAATCCAGAGAT	663
Db	2151	AGCTTCGAAACACCC	TGATTCACGTGCATATTCTCGCGCTGGAATGTTTCCGAG	2210
QY	664	CTTCCTCATGAGCGCTG	CCCTCCATGCGCAGCCCTCTGCAATTCTATGTGTGAACAGT	723
Db	2211	CTTGAAAACATGGCTCT	TCCCCCTTGTCACTTGCTTTTCAGCTCATGTATGTCCGATGGC	2270
QY	724	GAGCTGTCTGCACGCTGT	TACGAGATCGGGAGACATGGGCTTCGTGTGCTTTCAAC	783
Db	2271	AAACTGTCTTGGCACTCT	TACAGCGTTTCGCGAGATGTTCTCGGATGTGCTTTCAAC	2330
QY	784	ATACGACGTTACGCGCTCT	CTCAGTACGATGTGAGCATTCAGGGCTCGAAGCCAGAT	843
Db	2331	ATGCGATCTTATGCACTG	CTACCCACATGTTTGCACAGGAGAGGCTTGAAGTCGCG	2390
QY	844	GACTTTATACACATTTT	GGAGATGACATATTTTACTGAATCACTGAGCCACTGA	903
Db	2391	GAGTTCAATTGGACGTG	CGGGGACGTGCACATTTTATGACAACACAGAGAACGTCGCG	2450
QY	904	ATTACACTTCAGCGAGAA	CCCGACCTTCCCAAGCTCAGAGATTTCTTGAAAAGTATG	963
Db	2451	GAGCACTATGCGCGAAG	CTCGGCTTCGCCCA--CCTTGAGCTCAACAGAGGAGGG	2507
QY	964	AAATATTGAGACTTTCAA	AGTGAAGATTTCAATTCGATCCCACT	1023
Db	2508	TTCATGTTTGAATACAG	TTCGATGATCATCAGGTCCGCTAGATTCACACCATTTG	2567
QY	1024	AT	1025	
Db	2568	AT	2569	

```

RESULT 12
US-09-867-550-395
; Sequence 395, Application US/09867550
; Patent No. US20020082206A1
GENERAL INFORMATION:
APPLICANT: Leach, Martin D.
APPLICANT: Mehraban, Foad,
APPLICANT: Conley, Pamela
APPLICANT: Law, Debbie
APPLICANT: Topper, James
TITLE OF INVENTION: NO. US20020082206A1el Polynucleotides from Atherogenic Cells and
FILE OF INVENTION: Theredy
FILE REFERENCE: 21402-013 (Cura-3.13)
CURRENT APPLICATION NUMBER: US/09/867,550
CURRENT FILING DATE: 2001-09-20
PRIOR APPLICATION NUMBER: USSN 60/208,427
PRIOR FILING DATE: 2000-05-30
NUMBER OF SEQ ID NOS: 2125
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 395
LENGTH: 408
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)..(2)
OTHER INFORMATION: wherein any n is one of a or t or g or c
US-09-867-550-395

```

	Query Match	7.7%;	Score 118.8;	DB 10;	Length 408;
	Best Local Similarity	56.3%;	Pred. No. 1.9e-22;		
	Matches 222;	Conservative 0;	Mismatches 172;	Indels 0;	Gaps 0;
QY	524	GGCATTTTGGGCGCATACAGAGATATGATATCATGATTTTCAAGACGAGGGAGTTGACC	583		
Db	14	GGCGTATTTTCGGCGCACTAAATAATGCCGATGGTACTCTTAAACCAAGATGGTTTGAAC	73		
QY	584	AATCGAAGAGATGATTGACACCATCAAAACCAACCTGACGACAGAAATCAATCATCGT	643		
Db	74	AAATTAAAGGTGTTAATTAGAAATTTAAACCAACCCAAATTCAGCCGCTTTAATATGCT	133		

QY	644	GCCTTGGAAATCCAGAGATCTTCTCTGATAGGCGCTGCTTCATGCAATGCCCTCTGCG	703
Db	134	CTGTGTGGAAATCCAAATGAGGACAGACAAATTGATCAATACCGCTTGCATACATCTTCTTC	199
QY	704	AGTTCTATGTGTGAAACAGTACGATGTCCTTGCAGCTGTACACAGATCGGAGACATGG	763
Db	194	AATTTTGTGTCAAGACACAAAGTTGTTCATGCGACAGCTTATCAACGTAGCGCTGACGTAT	253
QY	764	GCCTGGGTGGCTTTCAACTACCGACGTACGCGCTGTACGTACATGATTTGGGACA	823
Db	254	TCTTAGGGGTGCCATTTAATATATGACGCTATGCGCTCTACCCAAATGATTTGTCAAG	313
QY	824	TCAACGGGCTGAACAGGTGATCTTTATACACTTTTGGAGATGCACTATTTTACGTA	883
Db	314	TGTGTGGCTTAGGTGTGTGTATTTTGTTTGACAGGTGGGATACGATCTTTTACGAA	373
QY	884	ATCAGATCGAGCCACTGAAATTTCAAGTTCAGGG	917
Db	374	ACCATTTTGAACAGGCAAAATCTACAAATTTGACGGG	407

```

RESULT 13
US-09-974-300-5027
Sequence 5027, Application US/09974300
Patent No. US2002014672A1
GENERAL INFORMATION:
APPLICANT: Berka, Randy M.
APPLICANT: Clausen, Ib Groh
TITLE OF INVENTION: Methods For Monitoring Multiple Gene
TITLE OF INVENTION: Expression
FILE REFERENCE: 10085,500-US
CURRENT APPLICATION NUMBER: US/09/974,300
CURRENT FILING DATE: 2001-10-05
PRIORITY APPLICATION NUMBER: 09/660,598
PRIORITY FILING DATE: 2000-10-06
PRIORITY APPLICATION NUMBER: 60/279,526
PRIORITY FILING DATE: 2001-03-27
NUMBER OF SEQ ID NOS: 8481
SOFTWARE: FastSeq For Windows Version 4.0
SEQ ID NO 5027
LENGTH: 213
TYPE: DNA
ORGANISM: Bacillus clausii
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)...(213)
OTHER INFORMATION: n = A,T,C or G
US-09-974-300-5027

```

	Query Match	4.4%	Score 68.2	DB 10	Length 213
	Best Local Similarity	57.3%	Pred. No. 5	2e-09	
	Matches 121	Conservative 0	Mismatches 90	Indels 0	Gaps 0
Qy	833	TGAAGCAGGTGACCTTATATACACTTTGGAGATGACATATTACTGTGATCAATCG	892		
Db	3	TAAAGCTGGGATTTATTCACAGTTGGAGATCCCATTTGTACAGCAACATATG	62		
Qy	893	AGGCATATAAATTCAGCTTCAGGAGAACCCAGACCTTTCCCAAGTCAGATCTTC	952		
Db	63	AACAGACAAAGCTTCATTTAAGCGGTAAACCAAAAGCTGCCAGATCGATTCATC	122		
Qy	953	GAAAGTTGAAAATATGANGACTTCAAAGCTGAAGACTTTGAGTTGAAGGTACATC	1012		
Db	123	CAGAGGTAAACGACCTGTTGGCTTTACTTACAGTGAATTTTGAACGTAAACATATGATC	182		
Qy	1013	CGCATTCACACTATTAAATATGAATAGAGCTGT	1043		
Db	183	CTCATCCGCAATTTAAAGCAGAGGGTCCGT	213		

RESULT 14
US-09-070-927A-315/c

OY	61	CCGCGCGGCACTGCTGCCTGCTCCGTCCTCCGCGCGCGCGCGCAATGCGCTGAGCGCGC	120
Db	61	CCCGCGCGCACTGCTGCCTGCTCCGTCCTCCGCGCGCGCGCGCAATGCGCTGAGCGCGC	120
OY	121	TGGAGGCTGCGCGCGCGCGCTTCGCTCCCGCGCGCAAGGAGGAGGAGCGCGAGCGCGCT	180
Db	121	TGGAGGCTGCGCGCGCGCGCTTCGCTCCCGCGCGCAAGGAGGAGGAGCGCGAGCGCGCT	180
OY	181	CCGCGCGAGGGGAGCTGCAAGTACTTGAGGAGAGATCACAACATCTCCGCTGCGCGCTC	240
Db	181	CCGCGCGAGGGGAGCTGCAAGTACTTGAGGAGAGATCACAACATCTCCGCTGCGCGCTC	240
OY	241	AGGAAGAGCAACGCAACGAGCAACGAGCAACCTGTGCGATATTCGCGATGCGAGCGCGCTAC	300
Db	241	AGGAAGAGCAACGCAACGAGCAACGAGCAACCTGTGCGATATTCGCGATGCGAGCGCGCTAC	300
OY	301	AGCGTGAAGATGAAATTCCTCTGCTGACACCAAACTGTGTTCGGAAGGGTGTTTTG	360
Db	301	AGCGTGAAGATGAAATTCCTCTGCTGACACCAAACTGTGTTCGGAAGGGTGTTTTG	360
OY	361	GAGGAGTTCCTGTGCTTTATCAAGGATCCACAAATGCTAAAGCTGTCTTCCAGAGA	420
Db	361	GAGGAGTTCCTGTGCTTTATCAAGGATCCACAAATGCTAAAGCTGTCTTCCAGAGA	420
OY	421	GTGAAATCTGCGGATCCCAATGATCCCGAGCTTTTGGACACGCTGAGATTCTCAC	480
Db	421	GTGAAATCTGCGGATCCCAATGATCCCGAGCTTTTGGACACGCTGAGATTCTCAC	480
OY	481	AGGAGAAGAGGGAGCTGCGGCGCGATTAATGCTTCAAGTGAAGCATTTTGCGGACAGA	540
Db	481	AGGAGAAGAGGGAGCTGCGGCGCGATTAATGCTTCAAGTGAAGCATTTTGCGGACAGA	540
OY	541	TACAGAGATATGAGATCAGATTAATTCAGGACAGGAGATTGACAACTGCAAGAGTGAAT	600
Db	541	TACAGAGATATGAGATCAGATTAATTCAGGACAGGAGATTGACAACTGCAAGAGTGAAT	600
OY	601	GACACCATCAAAACCAACCTTCAGACAGAGAAATCATATGTCGCTTGGAAATCAAGA	660
Db	601	GACACCATCAAAACCAACCTTCAGACAGAGAAATCATATGTCGCTTGGAAATCAAGA	660
OY	661	GATCTTCCTCTGATGCGCGCTGCTCCATGCTCATGCCCTTCGACATTTATGTGTGAAC	720
Db	661	GATCTTCCTCTGATGCGCGCTGCTCCATGCTCATGCCCTTCGACATTTATGTGTGAAC	720
OY	721	AGTGAAGCTTCGCGCGAGCTGTACCAAGATGCGGAGCAAGGAGCTGCGGTGCTCTTC	780
Db	721	AGTGAAGCTTCGCGCGAGCTGTACCAAGATGCGGAGCAAGGAGCTGCGGTGCTCTTC	780
OY	781	AACATTCGACAGTACGCTGCTGTACAGTAATGTGCGACATGACGCGGCTGAAGCA	840
Db	781	AACATTCGACAGTACGCTGCTGTACAGTAATGTGCGACATGACGCGGCTGAAGCA	840
OY	841	GGTGACTTTATACACATTTTGAGAGATGACATATTTACTGATCAATCAGACCACTG	900
Db	841	GGTGACTTTATACACATTTTGAGAGATGACATATTTACTGATCAATCAGACCACTG	900
OY	901	AAATTCAGCTTCAAGAGAACCCAGACCTTTCCAAAGCTCAGATTTCTCGAAAATT	960
Db	901	AAATTCAGCTTCAAGAGAACCCAGACCTTTCCAAAGCTCAGATTTCTCGAAAATT	960
OY	961	GAGAAATATGATGATTTCAAGCTTGAAGATTTGAGATGAGAGATGAGATGAGATGAGAT	1020
Db	961	GAGAAATATGATGATTTCAAGCTTGAAGATTTGAGATGAGAGATGAGATGAGATGAGAT	1020
OY	1021	ACTATTAATAAGAAATGCGCTTTAAGGGTCTTCAAGAGAGCTGAAAGATTTGTCA	1080
Db	1021	ACTATTAATAAGAAATGCGCTTTAAGGGTCTTCAAGAGAGCTGAAAGATTTGTCA	1080
OY	1081	GTCTTTAAGGGTGGGCTGAGTGCAGAGGTAAATTTCTTTTGTCTTAAAGAAAG	1140
Db	1081	GTCTTTAAGGGTGGGCTGAGTGCAGAGGTAAATTTCTTTTGTCTTAAAGAAAG	1140
OY	1141	AACATGCTCAAAATATGCTGCTGAGCTATCAGTTAATTTTAAAGAGTGTGCACT	1200

Db	1141	ACTGAGTCGCAAAAATCTGTCCGACCTTACGATTATTTTAAGGATGTCACCT	1200
Qy	1201	GGCAATGAACTGTGCGAGTTCTTTCATATATAAAGCTTGAATTACACTGAGG	1260
Db	1201	GGGAATGTAACTGTGCGAGTTCTTTCATATATAAAGCTTGAATTACACTGAGG	1260
Qy	1261	GTATGTGCAAAATGCTGAGGTTATGACAAAGTGAGAGATGAATGATGCTTTAG	1320
Db	1261	GTATGTGCAAAATGCTGAGGTTATGACAAAGTGAGAGATGAATGATGCTTTAG	1320
Qy	1321	CAAAAACATGATGTGATTTCAATCCAGCTACTTAAAGAAGGTGGTGAATTCAC	1380
Db	1321	CAAAAACATGATGTGATTTCAATCCAGCTACTTAAAGAAGGTGGTGAATTCAC	1380
Qy	1381	AAAGCTATTTTGAATATTTTAAATTTTAAAGATTTCAACAGCTATTCCTCAAT	1440
Db	1381	AAAGCTATTTTGAATATTTTAAATTTTAAAGATTTCAACAGCTATTCCTCAAT	1440
Qy	1441	CTGAGGAGCTGAAATACCAATCCATGATATGATGAGTGTATGAATTTAAGT	1500
Db	1441	CTGAGGAGCTGAAATACCAATCCATGATATGATGAGTGTATGAATTTAAGT	1500
Qy	1501	TGTTTATATGTGCTGCTATATAAAGAGTGTTGC	1536
Db	1501	TGTTTATATGTGCTGCTATATAAAGAGTGTTGC	1536

RESULT 3

US-09-347-878-29

Sequence 29, Application US/09347878C

Patent No. 6376210

GENERAL INFORMATION:

APPLICANT: Yuan, Chong

TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR ASSAYING ANALYTES

FILE REFERENCES: 25865-1651

CURRENT APPLICATION NUMBER: US/09/347,878C

CURRENT FILING DATE: 1999-07-06

NUMBER OF SEQ ID NOS: 75

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 29

LENGTH: 942

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION: Human thymidylate synthase gene: exons 1-8

PUBLICATION INFORMATION:

DATABASE ACCESSION NUMBER: D00596/GenBank

US-09-347-878-29

Query Match	61.3%; Score 942; DB 4; Length 942;		
Best Local Similarity	100.0%; Pred. No. 2.9e-241;		
Matches 942; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	106	ATGCTGTGCGCGGCTCGAGCTGCGCGCGCGGCTTGCCCGCGCGCGACGAGCGG	165
Db	1	ATGCTGTGCGCGGCTCGAGCTGCGCGCGGCTTGCCCGCGCGCGACGAGCGG	60
Qy	166	GAGCGCGAGCGCGCTCGCGCGCGCGAGAGCTCACTTCTGAGCAACGATCCACATC	225
Db	61	GAGCGCGAGCGCGCTCGCGCGCGCGAGAGCTCACTTCTGAGCAACGATCCACATC	120
Qy	226	CTCCGCTGCGCGCTCAAGAAAGACGACCGGACCGGACCGGACCTGTGCGATTCCG	285
Db	121	CTCCGCTGCGCGCTCAAGAAAGACGACCGGACCGGACCGGACCTGTGCGATTCCG	180
Qy	286	ATGCAAGCGCGCTACAGCTTGAAGATGAATTCCTCTGCTGACACCAACGTTGTTTC	345
Db	181	ATGCAAGCGCGCTACAGCTTGAAGATGAATTCCTCTGCTGACACCAACGTTGTTTC	240
Qy	346	TGAAAGGATGTTTGAAGAGTGTCTGTGTTTATCAAGGATCCACAAATGCTAAAGAG	405
Db	241	TGAAAGGATGTTTGAAGAGTGTGTGTATCAAGGATCCACAAATGCTAAAGAG	300

QY 406 CTGCTTCCAGAGGAGTGAATCTGGATGCAATGATCCGAGCTTTTGGACAGC 465
DB 301 CTGCTTCCAGAGGAGTGAATCTGGATGCAATGATCCGAGCTTTTGGACAGC 360
QY 466 CTGGAGTCTCCACCAAGAGAGAGAGGAGCTTGGCCCACTTTATGGCTTCCAGTGGAGG 525
DB 361 CTGGAGTCTCCACCAAGAGAGAGAGGAGCTTGGCCCACTTTATGGCTTCCAGTGGAGG 420
QY 526 CATTTTGGGGGAGATACAGAGATATGATCATTTTTCAGACAGAGAGTGGACCA 585
DB 421 CATTTTGGGGGAGATACAGAGATATGATCATTTTTCAGACAGAGAGTGGACCA 480
QY 586 CTGCAAGAGATTTGACACCATCAAAACCAACCTTACACAGAGAGATCATCATGTC 645
DB 481 CTGCAAGAGATTTGACACCATCAAAACCAACCTTACACAGAGAGATCATCATGTC 540
QY 646 GGTGGAGATCCAGAGATCTTCTGATGGGCGTGGCTCCAGCCATGCGCTTGGCAG 705
DB 541 GGTGGAGATCCAGAGATCTTCTGATGGGCGTGGCTCCAGCCATGCGCTTGGCAG 600
QY 706 TTCTATGTTGGAACAGTGAAGTGTCTCCAGCTGTACAGAGATGGAGAGATGGAGC 765
DB 601 TTCTATGTTGGAACAGTGAAGTGTCTCCAGCTGTACAGAGATGGAGAGATGGAGC 660
QY 766 CTGGGTGTGCTTTCAACATGCGCAGCTAGCGCTGTACAGATCATGATTTGGCAGATC 825
DB 661 CTGGGTGTGCTTTCAACATGCGCAGCTAGCGCTGTACAGATCATGATTTGGCAGATC 720
QY 826 ACGGAGCTTGAAGCAGGTGACTTTATACACACTTTGGAGATGACATATTTTCTGAT 885
DB 721 ACGGAGCTTGAAGCAGGTGACTTTATACACACTTTGGAGATGACATATTTTCTGAT 780
QY 886 CACATGAGCCACTGAAATTTAGCTTGAAGAGAACCCACACTTTTCCAAAGCTCAGG 945
DB 781 CACATGAGCCACTGAAATTTAGCTTGAAGAGAACCCACACTTTTCCAAAGCTCAGG 840
QY 946 ATTTCTGAAAAGTTGAGAAAATTTGATGATCAAGCTGAAGACTTTGATTTGAAGG 1005
DB 841 ATTTCTGAAAAGTTGAGAAAATTTGATGATCAAGCTGAAGACTTTGATTTGAAGG 900
QY 1006 TACATCCGATCCACTATTTAAATGAAATGCGTGTAG 1047
DB 901 TACATCCGATCCACTATTTAAATGAAATGCGTGTAG 942

RESULT 4

US-08-965-048-5/c
; Sequence 5, Application US/08965048
; Patent No. 6323244
; GENERAL INFORMATION:
; APPLICANT: Chen, Hong
; APPLICANT: Freimer, Nelson
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF NEUROPSYCHIATRIC DISORDERS
; FILE REFERENCE: 7853-093
; CURRENT APPLICATION NUMBER: US/08/965,048
; CURRENT FILING DATE: 1997-11-05
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 45716
; TYPE: DNA
; ORGANISM: Homo sapiens
US-08-965-048-5

Query Match 40.9%; Score 627.8; DB 4; Length 45716;
Best Local Similarity 91.1%; Pred. No. 9,4e-157;
Matches 665; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 807 GTACATGATTTGGCAGATCAAGGCTGAAGCCAGGTGACTTTATACACACTTTGGAGAG 866
DB 44683 GAACCTTGTATGATCATCCGTGTACTTGTTCATGAGACATGAGGCAATTACAACAG 44624

QY 867 TGGACATATTTTACTGATACATGAGCCACTGAAATTTCAAGCTTCAGGAGAACCCAG 926
DB 44623 GTGATACATATTTAGCCAAATATATGAGCTTATTTTGTATTTTACCTTCAGGAGAACCCAG 44564
QY 927 ACCCTTCCCAAGGTCAGGATCTTCGAAAAGTTGAGAAAATTTGATGACTTCAAGACTGA 986
DB 44563 ACCCTTCCCAAGGTCAGGATCTTCGAAAAGTTGAGAAAATTTGATGACTTCAAGACTGA 44504
QY 987 AGACTTCAGATTTAGAGGATCAATCCGATCCACTATTTAAATGAAATGAGCTGTTTA 1046
DB 44503 AGACTTCAGATTTAGAGGATCAATCCGATCCACTATTTAAATGAAATGAGCTGTTTA 44444
QY 1047 GGGTCTTTCAAAGAGCTTGAAGATATTTGACATTTTGAAGGCTTGGCTGGATGCCG 1106
DB 44443 GGGTCTTTCAAAGAGCTTGAAGATATTTGACATTTTGAAGGCTTGGCTGGATGCCG 44384
QY 1107 AGGTAAAGTTCTTTTGTCTTAAAGAAAGAACTAGGTCAAAAATCTGCTGTAC 1166
DB 44383 AGGTAAAGTTCTTTTGTCTTAAAGAAAGAACTAGGTCAAAAATCTGCTGTAC 44324
QY 1167 CTATCAGTTATTTATTTTGAAGATTTGCACTGGCAATGTAATGTCAGTTCCTT 1226
DB 44323 CTATCAGTTATTTATTTTGAAGATTTGCACTGGCAATGTAATGTCAGTTCCTT 44264
QY 1227 CCTATTTAAAGGCTTTGATTAATCTCACTGAGGCTATCTGACATGCTGAGTTATGA 1286
DB 44263 CCTATTTAAAGGCTTTGATTAATCTCACTGAGGCTATCTGACATGCTGAGTTATGA 44204
QY 1287 CAAAGTGAAGAGATGAATGATGCTTTAGCAAAAACATGATGTCATTTCAATC 1346
DB 44203 CAAAGTGAAGAGATGAATGATGCTTTAGCAAAAACATGATGTCATTTCAATC 44144
QY 1347 CCAAGTACTTATTAAGAAAGTTGATGATTTCAAGCTATTTTGAATTTTGA 1406
DB 44143 CCAAGTACTTATTAAGAAAGTTGATGATTTCAAGCTATTTTGAATTTTGA 44084
QY 1407 TATTTAAGATTTCAAGCTATTTCCCTCAATTTGAGGAGGCTGAGTACACATCCA 1466
DB 44083 TATTTAAGATTTCAAGCTATTTCCCTCAATTTGAGGAGGCTGAGTACACATCCA 44024
QY 1467 TCATGATGAGTGTGTATGAATTTANAGTTTNTANAGTGTATTAATAA 1526
DB 44023 TCATGATGAGTGTGTATGAATTTANAGTTTNTANAGTGTATTAATAA 43964
QY 1527 AGTGTCTGC 1536
DB 43963 AGTGTCTGC 43954

RESULT 5

US-08-965-048-6/c
; Sequence 6, Application US/08965048
; Patent No. 6323244
; GENERAL INFORMATION:
; APPLICANT: Chen, Hong
; APPLICANT: Freimer, Nelson
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF NEUROPSYCHIATRIC DISORDERS
; FILE REFERENCE: 7853-093
; CURRENT APPLICATION NUMBER: US/08/965,048
; CURRENT FILING DATE: 1997-11-05
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 45983
; TYPE: DNA
; ORGANISM: Homo sapiens
US-08-965-048-6

Query Match 40.9%; Score 627.8; DB 4; Length 45983;
Best Local Similarity 91.1%; Pred. No. 9,5e-157;
Matches 665; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

```

QY 807 GTACATGATGGGACATCAAGGAGCTGAGGAGCTTTATACACTTTGGGAGA 866
Db 44956 GAACCTTTGATGACATCCTGATCTGTTTCAATGACATGAGAGCAATTAACAG 44897
QY 867 TGCACATATTTTACCTGATCAATCAGACCACTGAAAATTCAGCTTCAGCGAACCAG 926
Db 44896 GTGCTACATTTTGGCAAAATTAATGAGCTTATTTTGTGTTTACGCTTACGAGAACAG 44837
QY 927 ACCCTTCCAAAGCTCAGATCTTCCGAAAGTTGAGAAAATGATGACTTCAAAAGTGA 986
Db 44836 ACCCTTCCAAAGCTCAGATCTTCCGAAAGTTGAGAAAATGATGACTTCAAAAGTGA 44777
QY 987 AGACTTTGATGAGAGGATCAATCCGATCAAATTTTAAATGAGAAATGGCTGTTA 1046
Db 44776 AGACTTTGATGAGAGGATCAATCCGATCAAATTTTAAATGAGAAATGGCTGTTA 44717
QY 1047 GGGTCTTCAAGAGGCTGAGAGATTTGACCTTTAGGGGTTGGGCTGAGATGCG 1106
Db 44716 GGGTCTTCAAGAGGCTGAGAGATTTGACCTTTAGGGGTTGGGCTGAGATGCG 44657
QY 1107 AGGTAAAGTTCTTTTGTCTTAAAGAAAGAAAGAACTAGGTCAAATCTGCTGAC 1166
Db 44656 AGGTAAAGTTCTTTTGTCTTAAAGAAAGAAAGAACTAGGTCAAATCTGCTGAC 44597
QY 1167 CTATCAGTATTAATTTTAAAGATGTCACATGGCAATGTATCTGTCAGTTCTTT 1226
Db 44596 CTATCAGTATTAATTTTAAAGATGTCACATGGCAATGTATCTGTCAGTTCTTT 44537
QY 1227 CCATTAATTAAGGCTTTGATTAATCTGACAGGATCTGACAAATGCTGAGTTATGA 1286
Db 44536 CCATTAATTAAGGCTTTGATTAATCTGACAGGATCTGACAAATGCTGAGTTATGA 44477
QY 1287 CAAAGTGAAGAAATGAATGATGCTTGAAGAAATGATGATGATGATGATGATGATG 1346
Db 44476 CAAAGTGAAGAAATGAATGATGCTTGAAGAAATGATGATGATGATGATGATGATG 44417
QY 1347 CCAAGTGAAGAAATGAATGATGCTTGAAGAAATGATGATGATGATGATGATGATG 1406
Db 44416 CCAAGTGAAGAAATGAATGATGCTTGAAGAAATGATGATGATGATGATGATGATG 44357
QY 1407 TATTTAAGATTTCAAGATCTTCCCTCAATCTGAGGAGCTGATGAACCATGA 1466
Db 44356 TATTTAAGATTTCAAGATCTTCCCTCAATCTGAGGAGCTGATGAACCATGA 44297
QY 1467 TCATGATGAGAGTGTGTTATGAATTTANAGTTGTTTATGTTGCTATTAATGA 1526
Db 44296 TCATGATGAGAGTGTGTTATGAATTTANAGTTGTTTATGTTGCTATTAATGA 44237
QY 1527 AGTGTCTGTC 1536
Db 44236 AGTGTCTGTC 44227

```

RESULT 6

US-09-318-448-11

Sequence 11, Application US/09318448

Patent No. 6210350

GENERAL INFORMATION:

APPLICANT: Johnson, William G.

APPLICANT: Stencos, Edward S.

TITLE OF INVENTION: METHODS FOR DIAGNOSING, PREVENTING, AND TREATING

FILE REFERENCE: 601-1-057

CURRENT APPLICATION NUMBER: US/09/318,448

CURRENT FILING DATE: 1999-05-25

NUMBER OF SEQ ID NOS: 46

SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 11

LENGTH: 18596

TYPE: DNA

ORGANISM: Homo sapiens

US-09-318-448-11

```

Query Match 39.8%; Score 611.8; DB 4; Length 18596;
Best Local Similarity 90.4%; Pred. No. 1,1e-152;
Matches 665; Conservative 0; Mismatches 65; Indels 6; Gaps 1;

QY 807 GTACATGATGGGACATCAAGGAGCTGAGGAGCTTTATACACTTTGGGAGA 866
Db 15510 GAACCTTTGATGACATCCTGATCTGTTTCAATGACATGAGAGCAATTAACAG 15569
QY 867 TGCACATATTTTACCTGATCAATCAGACCACTGAAAATTCAGCTTCAGCGAACCAG 926
Db 15570 GTGCTACATTTTGGCAAAATTAATGAGCTTATTTTGTGTTTACGCTTACGAGAACAG 15629
QY 927 ACCCTTCCAAAGCTCAGATCTTCCGAAAGTTGAGAAAATGATGACTTCAAAAGTGA 986
Db 15630 ACCCTTCCAAAGCTCAGATCTTCCGAAAGTTGAGAAAATGATGACTTCAAAAGTGA 15689
QY 987 AGACTTTGATGAGAGGATCAATCCGATCAAATTTTAAATGAGAAATGGCTGTTA 1046
Db 15690 AGACTTTGATGAGAGGATCAATCCGATCAAATTTTAAATGAGAAATGGCTGTTA 15749
QY 1047 GGGTCTTCAAGAGGCTGAGAGATTTGACCTTTAGGGGTTGGGCTGAGATGCG 1106
Db 15750 GGGTCTTCAAGAGGCTGAGAGATTTGACCTTTAGGGGTTGGGCTGAGATGCG 15809
QY 1107 AGGTAAAGTTCTTTTGTCTTAAAGAAAGAAAGAACTAGGTCAAATCTGCTGAC 1166
Db 15810 AGGTAAAGTTCTTTTGTCTTAAAGAAAGAAAGAACTAGGTCAAATCTGCTGAC 15869
QY 1167 CTATCAGTATTAATTTTAAAGATGTCACATGGCAATGTATCTGTCAGTTCTTT 1226
Db 15870 CTATCAGTATTAATTTTAAAGATGTCACATGGCAATGTATCTGTCAGTTCTTT 15929
QY 1227 CCATTAATTAAGGCTTTGATTAATCTGACAGGATCTGACAAATGCTGAGTTATGA 1286
Db 15930 CCATTAATTAAGGCTTTGATTAATCTGACAGGATCTGACAAATGCTGAGTTATGA 15989
QY 1287 CAAAGTGAAGAAATGAATGATGCTTGAAGAAATGATGATGATGATGATGATGATG 1346
Db 15990 CAAAGTGAAGAAATGAATGATGCTTGAAGAAATGATGATGATGATGATGATGATG 16049
QY 1347 CCAAGTGAAGAAATGAATGATGCTTGAAGAAATGATGATGATGATGATGATGATG 1406
Db 16050 CCAAGTGAAGAAATGAATGATGCTTGAAGAAATGATGATGATGATGATGATGATG 16109
QY 1407 TATTTAAGATTTCAAGATCTTCCCTCAATCTGAGGAGCTGATGAACCATGA 1466
Db 16110 TATTTAAGATTTCAAGATCTTCCCTCAATCTGAGGAGCTGATGAACCATGA 16169
QY 1467 TCATGATGAGAGTGTGTTATGAATTTANAGTTGTTTATGTTGCTATTAATGA 1520
Db 16170 TCATGATGAGAGTGTGTTATGAATTTANAGTTGTTTATGTTGCTATTAATGA 16229
QY 1521 TAAAGAGTGTCTGC 1536
Db 16230 TAAAGAGTGTCTGC 16245

```

RESULT 7

US-09-230-637-1

Sequence 1, Application US/09230637

Patent No. 6264958

GENERAL INFORMATION:

APPLICANT: Hayward, Gary

APPLICANT: Nicholas, John

APPLICANT: Hardwick, J. Marie

APPLICANT: Reitz, Marvin

TITLE OF INVENTION: No. 6264958el Genes of Kaposi's Sarcoma

FILE REFERENCE: 1107.78372

CURRENT APPLICATION NUMBER: US/09/230,637

CURRENT FILING DATE: 1999-11-23

PRIOR APPLICATION NUMBER: 60/022,591

; PRIOR FILING DATE: 1996-07-25
 ; PRIOR APPLICATION NUMBER: PCT US 97/12931
 ; PRIOR FILING DATE: 1997-07-24
 ; NUMBER OF SEQ ID NOS: 62
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1
 ; LENGTH: 1014
 ; TYPE: DNA
 ; ORGANISM: Kaposi's sarcoma-associated herpesvirus
 US-09-230-637-1

Query Match 25.2%; Score 387.4; DB 4; Length 1014;
 Best Local Similarity 63.5%; Pred. No. 1.3e-93;
 Matches 592; Conservative 0; Mismatches 341; Indels 0; Gaps 0;

```

QY 115 GCCGCTGGAGCTGCCGCCGCCCTTGCCCCCGCAGAGAGCGGAGCCGAG 174
DB 82 GGCCTCTTGCTTTGGCCCCCGAGGTTGACCTTGTTGATTGACATGAGTTACTGG 141
QY 175 CCGCTCCGCGCAGCGGAGCTGCACTCTGCGGAGATCCACACATCTCCGCTGC 234
DB 142 GCTGAGAGCCCCCAGAGACTTCTGATCTCAGGCAATTGGAGAAATTTTGCCGT 201
QY 235 GGCCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 294
DB 202 GGCAGCAGTCACTGACCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 261
QY 295 GCGTACAGCTGAGAGATGATTCCTCTGCTGAGAACAAAGTGTCTTGGAAAGT 354
DB 262 AGGTATAGTCTGCGGAGCAGCTTCCCTACTAACCAAGAGCGGTGTGGAGAGC 321
QY 355 GTTTTGGAGAGCTGCTGTGTTTATCAGAGAGATCCAAATGCTAAGAGCTGTCC 414
DB 322 GTAGTGAAGAGCTGCTTTGTTTCTGAAGGAGAGTGTGCTCAGAGAGCTTTCAAG 381
QY 415 AAGGAGTGAATCTGGAGTGCAGATGATCCAGACTTTTGGACAGCTGGAGATTC 474
DB 382 AAGAGATCAAGATATGAGCAAAATGAGCTCAGAGAGATTCCTGAGCGGAGCTGC 441
QY 475 TCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 534
DB 442 GCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 501
QY 535 GCAGATACAGAGATATGATCAGATATATTCAGAGAGAGAGAGAGAGAGAGAG 594
DB 502 GCGGAGTACGTGAGCGGAGTGTGATATACAGAGAGAGAGAGAGAGAGAGAG 561
QY 595 GTGATTCACACATCAAAACCACTTGAACAGAGAGAGAGAGAGAGAGAGAGAG 654
DB 562 ATTGTGATTTAATATAAAATAATATCCAGAGATAGAGAGAGAGAGAGAGAG 621
QY 655 CCAAGAGATTTCTCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 714
DB 622 CCGAGGAGCTTGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 681
QY 715 GTGAACAGTGAAGTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 774
DB 682 GCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 741
QY 775 CTTTCAACATGCGCAGCTACGCTCAGCTCAGTATGATGAGAGAGAGAGAGAG 834
DB 742 CTTTAAATGCTGAGAGATATCTCTTAACTATATGAGAGAGAGAGAGAGAGAG 801
QY 835 AAGCAGAGTGAATTTATACACATTTGGAGATGACATATTTAATCTGATCAGAG 894
DB 802 AAGACCGGAGAGATTTATCAACGTTGGAGAGAGAGAGAGAGAGAGAGAGAGAG 861
QY 895 CCACTGAATTAATGAGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 954
DB 862 CCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 921
QY 955 AAGTTGAGAGAGATTTGATGATTTCAAGAGAGAGAGAGAGAGAGAGAGAGAG 1014
  
```

```

DB 922 TCTGTTCTTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 981
QY 1015 CATCCACTATTAATAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1047
DB 982 CATCCACTATTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1014
  
```

RESULT 8

US-08-770-379-17/C
 ; Sequence 17, Application US/08770379
 ; Patent No. 5849564

; GENERAL INFORMATION:
 ; APPLICANT: Chang, Yuan
 ; APPLICANT: Bohenzky, Roy A.
 ; APPLICANT: Russo, James J.
 ; APPLICANT: Edelman, Isidore S.
 ; APPLICANT: Moore, Patrick S.
 ; TITLE OF INVENTION: POLYPEPTIDES FROM KAPOSI'S SARCOMA-ASSOCIATED
 ; TITLE OF INVENTION: HERPESVIRUS, DNA ENCODING SAME AND USES THEREOF
 ; NUMBER OF SEQUENCES: 20
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSES: Cooper & Dunham LLP
 ; STREET: 1185 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 10036

; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/770,379

; FILING DATE:
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: White, John P.
 ; REGISTRATION NUMBER: 28,678
 ; REFERENCE/DOCKET NUMBER: 52342
 ; TELEPHONE: (212) 278-0400
 ; TELEFAX: (212) 391-0525

; INFORMATION FOR SEQ ID NO: 17:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 35100 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 US-08-770-379-17

Query Match 25.2%; Score 387.4; DB 2; Length 35100;
 Best Local Similarity 63.5%; Pred. No. 8.2e-93;
 Matches 592; Conservative 0; Mismatches 341; Indels 0; Gaps 0;

```

QY 115 GCCGCTGGAGCTGCCGCCGCCCTTGCCCCCGCAGAGAGCGGAGCCGAG 174
DB 21023 GGCCTCTTGCTTTGGCCCCCGAGGTTGACCTTGTTGATTGACATGAGTTACTGG 20964
QY 175 CCGCTCCGCGCAGCGGAGCTGCACTCTGCGGAGATCCACACATCTCCGCTGC 234
DB 20963 GCTGAGAGCGCCCAAGAGAACTTCAGATCTCAGGCAATTGGAGAAATTTTGCCGT 20904
QY 235 GGCCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 294
DB 20903 GGCAGCAGTCACTGACCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 20844
QY 295 GCGTACAGCTGAGAGATGATTCCTCTGCTGAGAACAAAGTGTCTTGGAAAGT 354
DB 20843 AGGTATAGTCTGCGGAGCAGCTTCCCTTAACTAACCAAGAGCGGTGTGGAGAGC 20784
QY 355 GTTTTGGAGAGTGTGTTTATCAGAGAGATCCAAATGCTAAGAGAGAGAGAGAG 414
  
```

```

Db 20783 GTAGTCAAGAGCTCTTGTGTTCTGAGAGGGAGAGTCTGACTCCAGAGGAGCTTTCAGAA 20724
OY 415 AAGGAGGTGAATCTGGGATGCGAATGATCCCAAGACTTTTGAACAGCTGGGATTC 474
Db 20723 ACAGGAGTCAAGATATGGGACAAAATGAGTCCAGAGAGTTCTCGAGCCGAGCCCTG 20664
OY 475 TCACACAGAGAAAGAGGAGCTTGGGAGCCAGTTTANAGCTCCAGTGAAGGCACTTTGGG 534
Db 20663 GCGCACAGAGAGGAGGAGGATTTGGGACCTGTTTACAGTTTCCAGTGAAGGCACTTTGGG 20604
OY 535 GCAGATPACAGAGATATGGAATTCAGATTATTCAGAGCAGGAGTTGACCACTGCAGAA 594
Db 20603 GCGGCGAGCTTGAGAGCGGAGTGTGACTATPACAGGCGAGGAGTTGACCAATTGTGTA 20544
OY 595 GTGATTGACACCATCAAAACCAACCTGACAGACAGAGAAATCATCATGTGCGCTTGGAT 654
Db 20543 ATTGTGATTTAATTAATAAATATACCGCAGATAGAGAAATCATTAATGTGTGAGAAC 20484
OY 655 CCAAGAGATCTTCTCTGATGAGCGCTGCTCCATGACATGATGAGTCCCTGCGCATTTATG 714
Db 20483 CCGGCGGACTGTGCTGTGATGAGCCCTCCGCTGCACTTGTATGTCAATTTATGTA 20424
OY 715 GTGAACAGTGAAGTCTCTGCTGCACTGTACCAAGATGCGAGACATAGGCTCCGCTG 774
Db 20423 GCTGACGAGTGAAGCTTCTCTGCTGCACTGTATACAGAGTGGAGACATAGGCTTGGAG 20364
OY 775 CTTTCAACATGCGCAGCTTACGCTGCTCCATGATACATGATTTGGGACATACAGGAGCT 834
Db 20363 CTTTTCACATTTGCGAGCTATTTCTCTTACCTTATATGCTTGTCTGATTTAGTGTCTT 20304
OY 835 AAGCAGGTGACCTTTATACACATTTGGGAGATGACATATTTACCTGAATCCATGAC 894
Db 20303 AGACCGGAGAGTTTATTCACAGTTGGAGATGCCCATCTPACAAAAGCATATAGAG 20244
OY 895 CCACGTGAATTTAGCTTCTGAGCGAGAACCCAGACTTTTCCAAAGCTCAGATTTCTG 954
Db 20243 CCACTACGCTGAGTGAAGCGGACCTCCAGCTCCCTTCCGCGCTGAGATATCCCG 20184
OY 955 AAGATTGAGAAATTTGATGATCTTCAAGCTGAGAGCTTTCAATGTAAGGATTCATCCG 1014
Db 20183 TCTGTTCTTCCATGAGAGATTTACACTGATGATTTTGAAGCTGTGAGCTACTGCCG 20124
OY 1015 CATCCACTATTAAATGGAATGCGCTTTAG 1047
Db 20123 CATCTACATTCGTAAGAGATGCGATTAG 20091

```

```

RESULT 9
US-08-757-6694-17/c
; Sequence 17, Application US/087576694
; Patent No. 6183751
; GENERAL INFORMATION:
; APPLICANT: Chang, Yuan
; APPLICANT: Bohenzky, Roy A.
; APPLICANT: Russo, James J.
; APPLICANT: Edelman, Isidore S.
; TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

```

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/757,6694
FILING DATE:
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 45185-F
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 35100 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-757-6694-17

Query Match 25.2%; Score 387.4; DB 4; Length 35100;
Best Local Similarity 63.5%; Pred. No. 8,2e-93;
Matches 592; Conservative 0; Mismatches 341; Indels 0; Gaps 0;

OY 115 GCGGCTGGAGCTCCGCGCGCCCTTGCCTCCCGCGCAGAGAGCGGAGCCGAG 174
Db 21023 GAGGCTTTGGCTTTGGCCCGGAGTTGACCTTGTGATGATGAGTAACTGAG 20964
OY 175 CCGGCTCCGCGCAGCGGAGAGTGTGAGTCTGGGAGAGTCCACATCTCCGCTGC 234
Db 20963 GCTGAGAGCGCCCGCAGAGACTTCAATGATCTGAGGAGATTTTGTGCGCT 20904
OY 235 GAGCTCAAGAAAGACGACCGCAGCGGACCGGACCTGTGGTATTCGACATGAGGCG 294
Db 20903 GAGAGATTCAGCTGACCGGACCGGACATAGGCACTCTCTATTTGACATGAGGCGC 20844
OY 295 CGCTPAGCTGAGAGATGATTTCTCTGCTGACACCAAGCTGTGTTTGAAGGT 354
Db 20843 AGGTATATGTCGCGGACCACTTCTCTTACACCAAGAGGAGTGTGTTGGAGAGC 20784
OY 355 GTTTGAGAGAGTGTGCTGCTTATCAAGAGATCCAAATGCTAAGAGCTGTCTCC 414
Db 20783 GTAGTCAAGAGCTCTTGTGTTCTGAGAGGAGTGTGACTCCAGAGAGCTTTCAGAA 20724
OY 415 AAGGAGGTGAATCTGGGATGCGAATGATCCCAAGACTTTTGAACAGCTGGGATTC 474
Db 20723 ACAGGAGTCAAGATATGGGACAAAATGAGTCCAGAGAGTTCTCGAGCCGAGCCCTG 20664
OY 475 TCACACAGAGAAAGAGGAGCTTGGGAGCCAGTTTANAGCTCCAGTGAAGGCACTTTGGG 534
Db 20663 GCGCACAGAGAGGAGGAGGATTTGGGACCTGTTTACAGTTTCCAGTGAAGGCACTTTGGG 20604
OY 535 GCAGATPACAGAGATATGGAATTCAGATTATTCAGAGCAGGAGTTGACCACTGCAGAA 594
Db 20603 GCGGCGAGCTTGAGAGCGGAGTGTGACTATPACAGGCGAGGAGTTGACCAATTGTGTA 20544
OY 595 GTGATTGACACCATCAAAACCAACCTGACAGACAGAGAAATCATCATGTGCGCTTGGAT 654
Db 20543 ATTGTGATTTAATTAATAAATATACCGCAGATAGAGAAATCATTAATGTGTGAGAAC 20484
OY 655 CCAAGAGATCTTCTCTGATGAGCGCTGCTCCATGACATGATGAGTCCCTGCGCATTTATG 714
Db 20483 CCGGCGGACTGTGCTGTGATGAGCCCTCCGCTGCACTTGTATGTCAATTTATGTA 20424
OY 715 GTGAACAGTGAAGTCTCTGCTGCACTGTACCAAGATGCGAGACATAGGCTCCGCTG 774
Db 20423 GCTGACGAGTGAAGCTTCTCTGCTGCACTGTATACAGAGTGGAGACATAGGCTTGGAG 20364
OY 775 CTTTCAACATGCGCAGCTTACGCTGCTCCATGATACATGATTTGGGACATACAGGAGCT 834
Db 20363 CTTTTCACATTTGCGAGCTATTTCTCTTACCTTATATGCTTGTCTGATTTAGTGTCTT 20304
OY 835 AAGCAGGTGACCTTTATACACATTTGGGAGATGACATATTTACCTGAATCCATGAC 894

```

QY	115	GCCTGCTGGAGCTCCGCGCCGGCCCTTGCCCGCCGACAGGAGGGGACGCTCGAG	174
Dh	21023	GGGTCTTTGGCTTTGGCCCCCGAGGTGACCTCTTGTGATTCAGATGATGATTAACGTAGG	20964
QY	175	CCGCTCCGCGCGCAGGGGAGCTGGCAGTACCTGGGCGAGATTCACACATCTCCGCTG	234
Dh	20963	GCTGAGAGCCCCACGGGAACTTCAGATCTCAGGACGTTGAGGGAATTTGTGCCGT	20904
QY	235	GGGTTCAGAAAGACGACCGCACGCGCACCCCTGTCCGATTCGCGANTCGAGCG	294
Dh	20903	GGAGAGATGCACTCGACCGCACCGGCACATAGGCACTCTCTATTTGGATCGAGGC	20844
QY	295	CGTACAGCTCGAGAGATGATTCCTCTGCTGACACCAACGTGTGTTCTGGAAAGGT	354
Dh	20843	AGGTATATCTCGCGGACCACTTCCCTTACACCAAGAGCGGTGTGTTGCGAGGC	20784
QY	355	GTTTTGGAGAGCTGTGTGTGTTTCAAGGAGTCCAAATGCTAAGAGCTGCTTCC	414
Dh	20783	GTATGTCAAGAGCTGTGTGTGTTTCAAGGAGGATCTGACTCCAGGGAGTTTCAAGA	20724
QY	415	AAGGAGTGAATAATCTGCAATGCCAATGGATCCGACACTTTTGGACAGCTCGGATTC	474
Dh	20723	ACGAGATTCAGATATTTGGACAAATAATGGCTTCAGGAGATTTCTGTGCGCGGGGCTG	20664
QY	475	TCACCCAGAGAAGAGGGGACCTTGGGCCCGAGTTTATGCTTCCATGTAGAGCATTTTGG	534
Dh	20663	GGGCACAAAGGAGGGGAGTTTGGGACCTGTTTACGTTTTCAGTGAAGGACACTTGG	20604

Dy 180 TCGCCGACGGGAGCTGCAGTACTG3GGCAGATCCAACACATCTCGCTGGCGGT 239
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 963 TCGGGGCCACGAGAAATTCCAGTACTCTGCACTCTATTGGCATTTATTCACAACGGAGC 1022

NAME: MONROY, Gladys H
 REGISTRATION NUMBER: 32,430
 REFERENCE/DOCKET NUMBER: 27340-20021.00
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 650-813-5600
 TELEFAX: 650-494-0792
 TELEX: 706141
 INFORMATION FOR SEQ ID NO: 1092:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 7685 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: circular
 MOLECULE TYPE: DNA (genomic)
 HYPOTHEetical: NO
 ANTI-SENSE: UNKNOWN
 ORIGINAL SOURCE:
 ORGANISM: PORPHYROMONAS GINGIVALIS
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: 1...7685
 US-09-221-017B-1092

Query Match 10.4%; Score 160.2; DB 4; Length 7685;
 Best Local Similarity 52.7%; Pred. No. 1,2e-32;
 Matches 443; Conservative 0; Mismatches 338; Indels 60; Gaps 2;

186 GCACGGGAGCTGCATACCTGGGGAGATCCACACATCCCGCTGGCGGTACAGAA 245
 4571 GCGTGTATGAAACAGTATTGGATCTCTTCGCCGATCTGAGAGAGGATACAAA 4512
 246 GGAAGACCGGACGGGACCGACCCCTGCGTATTCGGCATGCGGCGTACAGCT 305
 4511 AGAAGACCGGACGGGATCGGATGATGATGATGATGATGATGATGATGATGAT 4452
 306 GAGAGATGAAATTCCTCTCTCTGACCAACCAACGTTGTTCTGAAAGGTTTGAAGA 365
 4451 TGAAGAGGATTCCTCTCTCTGACCAACCAACGTTGTTCTGAAAGGTTTGAAGA 4392
 366 GTTGTGTGTTTATCAAGGATCCAAATGCTAAAGGTTCTTCCAAAGGATGAA 425
 4391 ACTGTTGTGTTCTCTCAAGGATGACAGCAAGTGCATTACCTGACAGAAATGAGTACG 4332
 426 AATCTGGGATGCAATGATCCGAGACTTTTGGAGAGCTGGGATTTCCACAGAGA 485
 4331 CAITTTGAACG-----AAATGGCTGATGA 4308
 486 AGAAGGAGACTTGGGCGCCAGTTTATGCTTCAGTGGAGGATTTTGGGCGAATACAG 545
 4307 GAAAGGAGATTTGGGACCTATATGCTTACCAAGTGGCGAGTTGGCG-----4259
 546 AGATATGAAATCAAGATTATTCAGACAGAGGATTGACCAATGCAAAAGGATGATGACAC 605
 4258 -----GACTATGCGGAGGAGGATATGACCAATGCAAAAGGATGATGACAC 4212
 606 CATCAAAACCAACCTTACAGACAGAAATATATATGAGGCTTGGAAATCCAGAGATCT 665
 4211 GGTGGCGAAGAACCGGACCGCGGATATGATGCTGCGAATGATGAGGCTGAGGCT 4152
 666 TCTCTGATGAGGCTGCTTCATGCTTCCATGCTTCCATGCTTCCATGCTTCCATGCT 725
 4151 TCCCATATGATGATGCTTCCATGCTTCCATGCTTCCATGCTTCCATGCTTCCATGCT 4092
 726 GGTGCTTCCAGCTGATCAAGAGATCGGAGACATGAGGCTGAGTCTCTTCAACAT 785
 4091 GGTGAGCTGCACTATATGAGGATGCGGATTTCTTCTGGGATACCTGTTAATAT 4032
 786 CGCCAGCTAGCCCTGCTCACTATGATGCGACATCAAGGCTTGAAGCGAGTGA 845
 4031 AGCTTCTGATGCACTCTCTGATCAATATGATGCTCAATGATCAAGGCTGAAACCGGCG 3972
 846 CTTATATACACTTTGGGAGATGCAATATTTACCTGATCAATGAGGCTGAAAT 905

Db 3971 ATTGTACATACATGCGGAGCGCCCATATCTATCTTAAACCTCCAGAGGTACAGCT 3912
 Qy 906 TCGGCTTACGCGAGAACCCAGACCTTCCCAAGCTCAGGATTTCTTGAAAGTTGAGAA 965
 Db 3911 ACAACTGATGCTGAGCCGAGAGGCTTGGCCCAATGATGATGATGATGATGATGAT 3852
 Qy 966 AATGATGACTTCAAGCTGAGAGCTTTCAGATTGAAAGGTTACATTCGATCCAACTAT 1025
 Db 3851 TCTTTCGACTTCGATACAGAGACTTCCGATCGAAGGATGATGATCGGATCCGACAT 3792
 Qy 1026 T 1026
 Db 3791 T 3791

RESULT 14
 US-09-203-895-1
 ; Sequence 1, Application US/09203895
 ; Patent No. 6436410
 ; GENERAL INFORMATION:
 ; APPLICANT: Krishnan, B. Rajendra
 ; APPLICANT: Yoder, S. Christine
 ; APPLICANT: Dursch, Becky A.
 ; TITLE OF INVENTION: DNA ENCODING NEOSPORA DIHYDROFOLATE REDUCTASE
 ; TITLE OF INVENTION: THYMIDYLATE SYNTHASE
 ; FILE REFERENCE: PC9690A
 ; CURRENT APPLICATION NUMBER: US/09/203,895
 ; EARLIER FILING DATE: 1998-12-02
 ; EARLIER APPLICATION NUMBER: 60/067,507
 ; EARLIER FILING DATE: 1997-12-04
 ; EARLIER APPLICATION NUMBER: 60/095,213
 ; NUMBER OF SEQ ID NOS: 9
 ; SOFTWARE: PatentIn Ver. 2.0 - beta
 ; SEQ ID NO 1
 ; LENGTH: 9603
 ; TYPE: DNA
 ; ORGANISM: Neospora caninum
 ; FEATURE:
 ; NAME/KEY: gene
 ; LOCATION: (2405)..(8199)
 US-09-203-895-1

Query Match 8.1%; Score 124.2; DB 4; Length 9603;
 Best Local Similarity 66.8%; Pred. No. 5e-23;
 Matches 193; Conservative 0; Mismatches 93; Indels 3; Gaps 1;

Qy 673 ATGCGGCTGCTTCATGCGATGCTCTTGCAGATTCTATATGCTGA---ACAATGAGCTG 729
 Db 7275 ATGCGGCTGCTTCATGCGATGCTCTTGCAGATTCTATATGCTGA---ACAATGAGCTG 7334
 Qy 730 TCTTCCAGCTGATACAGAGATGCGGAGATGCGGCTGCGGCTTCAACATGCGC 789
 Db 7335 TCTTCCAGCTGATACAGAGATGCGGAGATGCGGCTGCGGCTTCAACATGCGC 7394
 Qy 790 AGCTAGCGGCTGCTTCATGCGATGCTCTTGCAGATTCTATATGCTGA---ACAATGAGCTG 849
 Db 7395 TCTTCCAGCTGATACAGAGATGCGGAGATGCGGCTGCGGCTTCAACATGCGC 7454
 Qy 850 ATACACCTTGGGAGATGAGATATTTTACCTGATACATGAGCACTGAAATTCAG 909
 Db 7455 ATTCACCTTGGGAGATGAGATATTTTACCTGATACATGAGCACTGAAATTCAG 7514
 Qy 910 CTTTCAAGAGAACCCAGACCTTCCCAAGCTCAGGATTTCTTGAAAG 958
 Db 7515 CTGGCAGAGAACCCAGACCTTCCCAAGCTCAGGATTTCTTGAAAG 7563

RESULT 15
 US-08-714-918-7/c
 ; Sequence 7, Application US/08714918
 ; Patent No. 6037123
 ; GENERAL INFORMATION:

APPLICANT: Benton, Bret
APPLICANT: Lee, Ving
APPLICANT: Malouin, Francois
APPLICANT: Martin, Patrick K.
APPLICANT: Schmidt, Molly B.
APPLICANT: Sun, Dongxu
TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS ANTIBACTERIAL
TITLE OF INVENTION: TARGET GENES
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESS:
ADDRESSER: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: Storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/714,918
FILING DATE: September 13, 1996
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/009,102
FILING DATE: December 22, 1995
APPLICATION NUMBER: 60/003,798
FILING DATE: September 15, 1995
ATTORNEY/AGENT INFORMATION:
NAME: Wardburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 222/005
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 3479 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-714-918-7

Query Match 7.9%; Score 121.2; DB 3; Length 3479;
Best Local Similarity 56.8%; Pred. No. 1.9e-22;
Matches 268; Conservative 6; Mismatches 194; Indels 4; Gaps 3;

QY 578 TTGACCACTGCAAGAGTATGACACCAACCAACCCCTGACGACAGAGATCA 637
DB 1908 TTGATTAACCTTAAACAGTATGACAAATTAAGCATATTCGATTCGAGGACACA 1849
QY 638 TCATGCGCTTGAATCCAGAGATCTTCTCTGATGCGGCTGCTCATGCGATGCC 697
DB 1848 TCGTATCTGATGAAATCCAGAGAAATGATACATGSCATCTCGCTGTGATCCA 1789
QY 698 TCTGCCAGTCTTATGCTGATGACAGTGTCTTCCGACGCTGACGAGATCGGAG 757
DB 1788 TGTTCAGTTTATGTCAGAGTGTGATTAAGTTGCCAGTTATACCAAGTACGCGAG 1729
QY 758 ACATGCGGCTGCGTGCCTTCAACATGCGACATGCGCCGCTGCACTGACATGATG 817
DB 1728 ATATCTTTTAGTGTGCACTTATATGACAGCTACGCTTATGACACCTTATG 1669
QY 818 CGACATACG-GGCTGAAGCCAGGTGACTTATACACATTTGGAGATGACATATT 876
DB 1668 CCAAGAAATGTGACTTGAAGTGGGTATTTGTGATACATTTGAGATGACATATT 1609
QY 877 TACCTGATCAGATCGAGCCACTGAATTAAGCTTACGGGAGAACCAAGCTTTCCA 936

DB 1608 TATTCAATCATATTGATGCGATTCGMAACACATTAGAGCTGAAGCTTCATCTCC 1549
QY 937 AAGCTGAGATTCTTTCGAAAAGTTGAGAAAATTGATCTCAAGCTGAGACTTT-CA 995
DB 1548 AACTTA-AAAATTACAGGACAAAGTCTATTTCGACATTAATTAATGAAATTTGGA 1491
QY 996 GATTGAAGGTCATTCGCACTCCACTATTTAAATGAAATGCGCTGTTAG 1047
DB 1490 AATTGTGACTATGATCATCATCCAGCMATTAAGCTCMATAGCAATGTAG 1439

Search completed: November 28, 2002, 22:44:25
JOB Time : 209.177 secs

XX Human DNA sequence #200 expressed during foam cell differentiation.
 DE Human: foam cell differentiation; atherosclerosis; cerebral stroke;
 XX Human: foam cell differentiation; atherosclerosis; cerebral stroke;
 KM Cardiovascular disorder; coronary artery disease; gene therapy; ds.
 XX Hemo sapiens.
 OS Hemo sapiens.
 XX MO200177389-A2.
 PN 18-OCT-2001.
 PD 04-APR-2001; 2001MO-US11128.
 PF 05-APR-2000; 2000US-195106P.
 PR (INCYTE GENOMICS INC.
 PA Shiftman D, Somogyi R, Lawn R, Seilhamer JJ, Porter GJ, Mikita T,
 P1 Tai J;
 P1 WPI; 2002-010925/01.
 DR Composition useful for diagnosis of conditions, disorders or diseases
 XX associated with atherosclerosis, comprises several polynucleotides that
 XX are differentially expressed in foam cell development -
 XX Claim 1; Page 250-251, 315pp; English.

XX The present invention relates to the isolation of human polynucleotide
 CC sequences that are differentially expressed during foam cell
 CC differentiation. The polynucleotide sequences of the invention or a
 CC composition comprising these polynucleotides are useful as a high
 CC throughput method for detecting altered expression of one or more
 CC polynucleotides in a sample. The polynucleotides can be used in the
 CC diagnosis of disorders associated with foam cell development such as
 CC atherosclerosis, cerebral stroke, and cardiovascular disorders such as
 CC coronary artery disease. The polynucleotide sequences can also be used
 CC as PCR primers and probes. The polynucleotides of the invention are also
 CC useful in gene therapy. AAS94/46-AAS95021 represent the human
 CC polynucleotide sequences of the invention which are differentially
 CC expressed during foam cell differentiation.

SO Sequence 3298 BP; 898 A; 709 C; 744 G; 919 T; 28 other;

Query Match 97.2%; Score 1493.4; DB 24; Length 3298;

Best Local Similarity 99.3%; Pred. No. 0;

Matches 1510; Conservative 0; Mismatches 5; Indels 6; Gaps 1;

QY 22 CCGGCTCCGTCGCGCGCGCACTTGGCTGCTCCGCGCGCGCACTTGGCGT 81
 Db 1 CCGGCTCCGTCGCGCGCGCACTTGGCTGCTCCGCGCGCGCACTTGGCGT 60
 QY 82 CCGGCTCCGTCGCGCGCGCACTTGGCTGCTCCGCGCGCGCACTTGGCGT 141
 Db 61 CCGGCTCCGTCGCGCGCGCACTTGGCTGCTCCGCGCGCGCACTTGGCGT 120
 QY 142 TTGCCCCCGCGCGCGCGCACTTGGCTGCTCCGCGCGCGCACTTGGCGT 201
 Db 121 TTGCCCCCGCGCGCGCGCACTTGGCTGCTCCGCGCGCGCACTTGGCGT 180
 QY 202 TTGCCCCCGCGCGCGCGCACTTGGCTGCTCCGCGCGCGCACTTGGCGT 261
 Db 181 TTGCCCCCGCGCGCGCGCACTTGGCTGCTCCGCGCGCGCACTTGGCGT 240
 QY 262 ACCGCGACCTGTCGCTATTCGCGCGCGCACTTGGCTGCTCCGCGCGCGT 321
 Db 241 ACCGCGACCTGTCGCTATTCGCGCGCGCACTTGGCTGCTCCGCGCGCGT 300
 QY 322 TTGCTGACCAACCAAGTGTGTTTGAAGGAGTGTGTTGAAGGAGTGTGTTT 381
 Db 301 TTGCTGACCAACCAAGTGTGTTTGAAGGAGTGTGTTGAAGGAGTGTGTTT 360

QY 362 AAGGATCCACAAATGCTAAAGAGCTGTCTCCAAAGGAGTGAATCTGGATGCCAAT 441
 Db 361 AAGGATCCACAAATGCTAAAGAGCTGTCTCCAAAGGAGTGAATCTGGATGCCAAT 420
 QY 442 GGATCCGAGACTTTTGGACAGCTTGGAGTTCTCCACCAAGAGAGGAGCTTGGGCT 501
 Db 421 GGATCCGAGACTTTTGGACAGCTTGGAGTTCTCCACCAAGAGAGGAGCTTGGGCT 480
 QY 502 CCAATTATGAGCTTCAGTGAAGGAGCTTTGGGAGGAGTCAAGATATGAAATGAGT 561
 Db 481 CCAATTATGAGCTTCAGTGAAGGAGCTTTGGGAGGAGTCAAGATATGAAATGAGT 540
 QY 562 TATTCAGGACAGGAGTTCACCACTGCAAGAGTTCACCACTTGAATCAACCT 621
 Db 541 TATTCAGGACAGGAGTTCACCACTGCAAGAGTTCACCACTTGAATCAACCT 600
 QY 622 GAGGACAGAGATATCATGTGCGCTTGGAAATCCAGAGATCTTCTGATGCGCTG 681
 Db 601 GAGGACAGAGATATCATGTGCGCTTGGAAATCCAGAGATCTTCTGATGCGCTG 660
 QY 682 CTTCCAGGACATGCGCTTCCAGTTCATATGAGTGAACATGAGCTTCTGAGCTG 741
 Db 661 CTTCCAGGACATGCGCTTCCAGTTCATATGAGTGAACATGAGCTTCTGAGCTG 720
 QY 742 TACCAAGATCGGAGACATGAGGCTGCGCTTCAACATGCGCACTGAGGCTG 801
 Db 721 TACCAAGATCGGAGACATGAGGCTGCGCTTCAACATGCGCACTGAGGCTG 780
 QY 802 CTGACGATGATGATGCGCATGCGAGGCTGGAAGCGAGTGAATTAACACTTGG 861
 Db 781 CTGACGATGATGATGCGCATGCGAGGCTGGAAGCGAGTGAATTAACACTTGG 840
 QY 862 GAGGATGCAATATTTACCTGAATCAATGAGGCTGGAATTAACACTTGG 921
 Db 841 GAGGATGCAATATTTACCTGAATCAATGAGGCTGGAATTAACACTTGG 900
 QY 922 CCGAGACCTTTCCAAAGGCTCAGATTTCTGAAAGTGAAGTGAATTAACACTTGG 981
 Db 901 CCGAGACCTTTCCAAAGGCTCAGATTTCTGAAAGTGAAGTGAATTAACACTTGG 960
 QY 982 GCTGAAGACTTCAATTTGAAGGATCAATCCGATCCAACTTAATTAATGAAATGCT 1041
 Db 961 GCTGAAGACTTCAATTTGAAGGATCAATCCGATCCAACTTAATTAATGAAATGCT 1020
 QY 1042 GTTTAGGAGCTTTCAAGAGGCTTGAAGGATTTGCACTTTTGAAGGCTTGA 1101
 Db 1021 GTTTAGGAGCTTTCAAGAGGCTTGAAGGATTTGCACTTTTGAAGGCTTGA 1080
 QY 1102 TGCAGAGTAAAGTCTTTTCTCTAAAGAAAGAACTTGAAGTAAATGCTGCT 1161
 Db 1081 TGCAGAGTAAAGTCTTTTCTCTAAAGAAAGAACTTGAAGTAAATGCTGCT 1140
 QY 1162 GTGACCTATGATTTTAAAGGCTTTTGAAGGATTTGCACTTTTGAAGGCTTGA 1221
 Db 1141 GTGACCTATGATTTTAAAGGCTTTTGAAGGATTTGCACTTTTGAAGGCTTGA 1200
 QY 1222 TCTTTTCAATTAATAAGGCTTTTGAAGGATTTGCACTTTTGAAGGCTTGA 1281
 Db 1201 TCTTTTCAATTAATAAGGCTTTTGAAGGATTTGCACTTTTGAAGGCTTGA 1260
 QY 1282 ATGAACAAAGTGAAGGATTAATGATGCTTGAAGGATTAATGATGCTTGA 1341
 Db 1261 ATGAACAAAGTGAAGGATTAATGATGCTTGAAGGATTAATGATGCTTGA 1320
 QY 1342 CAATCCAGTACTTATTAAGAGGCTTGAAGGATTTTGAAGGATTTTGAAGTATTT 1401
 Db 1321 CAATCCAGTACTTATTAAGAGGCTTGAAGGATTTTGAAGGATTTTGAAGTATTT 1380
 QY 1402 TGAATATTTTAAAGATTTTCAAGCTATTTCCCAATTTGAAGGAGTGAATCAAC 1461
 Db 1381 TGAATATTTTAAAGATTTTCAAGCTATTTCCCAATTTGAAGGAGTGAATCAAC 1440
 QY 1462 ATGATCATGATGATGAGTGTGTTATGAATCTTTTAAAGTGTGTTATGCTG 1515

OY 1179 AATTCTTAAGGAT-GTTCACCTGCGAATGTAACCTGCGCAGTTCTTCCATATATAA 1237
 CC the disorder, by contacting a small molecule test compound (II) with an
 CC RTS protein (II) comprising a sequence of either 361 or 416 amino acids
 CC (AB05587 or AB05588) determining whether (I) binds to (II). (C) has
 CC antianemic, antidepressant, neuroleptic, nootropic and tranquilizer
 CC activities. RTS polynucleotide sequences can be used in gene therapy.
 CC The method is useful in identifying compounds for treating RTS-mediated
 CC neuropsychiatric disorders including bipolar affective disorder,
 CC e.g., severe bipolar affective (mood) disorder, bipolar affective (mood)
 CC disorder with hypomania, major depression, schizophrenia, attention
 CC deficit disorder and schizoaffective disorder. The present sequence
 CC represents human RTS-alpha gene which is used in the exemplification of
 CC the present invention. RTS has been located to chromosome 18.
 XX
 SQ Sequence 45716 BP; 12099 A; 10402 C; 10435 G; 12718 T; 62 other;
 Query Match 40.9%; Score 627.8; DB 24; Length 45716;
 Best Local Similarity 91.1%; Pred. No. 2.9e-153;
 Matches 665; Conservative 0; Mismatches 65; Indels 0; Gaps 0;
 OY 1476 AGAGTGTGTTTANGACT-----TTANAGTTGTTTATATANGTTGCTATATTAAGAGT 1529
 Db 159 AGAGTGTGTTATGAACTTAAAGTTATGTTTATATGTTGCTATATTAAGAGT 100
 OY 1530 GTTCTGC 1536
 Db 99 GTTCTGC 93
 RESULT 4
 ID ABA93401/c
 ID ABA93401 standard; DNA; 45716 BP.
 XX ABA93401;
 AC
 XX 22-APR-2002 (first entry)
 DT
 XX
 DB Human RTS-alpha gene SEQ ID NO:5.
 XX
 XX Human; RTS-alpha; RTS-beta; RTS; thymidylate synthase; chromosome 18;
 KW mutant thymidylate synthase; enzyme; antianemic; antidepressant;
 KW neuroleptic; nootropic; tranquilizer; gene therapy; hypomania;
 KW neuropsychiatric disorder; bipolar affective disorder; schizophrenia;
 KW severe bipolar affective disorder; bipolar affective disorder;
 KW major depression; attention deficit disorder; schizoaffective disorder;
 KW gene; ds.
 XX
 XX Homo sapiens.
 OS
 XX US6323244-B1.
 PN
 XX 27-NOV-2001.
 PD
 XX 05-NOV-1997; 97US-0965048.
 PF
 XX 05-NOV-1997; 97US-0965048.
 PR
 XX (MILL-) MILLENNIUM PHARM INC.
 PA (REGC) UNIV CALIFORNIA.
 PA
 XX
 XX Chen H, Frejmer NB;
 PI
 XX WPI; 2002-112960/15.
 DR
 XX
 XX Treating an RTS (mutant form of thymidylate synthase) - mediated
 PT neuropsychiatric disorder, e.g., bipolar affective disorder, comprises
 PT administering a compound, identified by an RTS assay, that interferes
 PT with RTS gene product -
 XX
 XX Disclosure; Fig 3A-S; 110pp; English.
 PS
 CC The present invention describes a method for treating an RTS (a mutant
 CC form of thymidylate synthase) mediated neuropsychiatric disorder. The

CC method comprises identifying a compound (C) that can be used to treat
 CC the disorder, by contacting a small molecule test compound (II) with an
 CC RTS protein (II) comprising a sequence of either 361 or 416 amino acids
 CC (AB05587 or AB05588) determining whether (I) binds to (II). (C) has
 CC antianemic, antidepressant, neuroleptic, nootropic and tranquilizer
 CC activities. RTS polynucleotide sequences can be used in gene therapy.
 CC The method is useful in identifying compounds for treating RTS-mediated
 CC neuropsychiatric disorders including bipolar affective disorder,
 CC e.g., severe bipolar affective (mood) disorder, bipolar affective (mood)
 CC disorder with hypomania, major depression, schizophrenia, attention
 CC deficit disorder and schizoaffective disorder. The present sequence
 CC represents human RTS-alpha gene which is used in the exemplification of
 CC the present invention. RTS has been located to chromosome 18.
 XX
 SQ Sequence 45716 BP; 12099 A; 10402 C; 10435 G; 12718 T; 62 other;
 Query Match 40.9%; Score 627.8; DB 24; Length 45716;
 Best Local Similarity 91.1%; Pred. No. 2.9e-153;
 Matches 665; Conservative 0; Mismatches 65; Indels 0; Gaps 0;
 OY 807 GTACATGATGGCCACATACAGGGGCTGAAAGCCAGGTGACTTATACACTTTGGGAGA 866
 Db 44683 GAACCTTGTGATCAATCCTGTGACTTGTTCATGACATGAGAGCAATTAACAAG 44624
 OY 867 TGACATATATTTACCTGAATCAGATGAGCCACTGAATTCAGCTTCAGAGAACCCAG 926
 Db 44623 GTGCTGAATTTATGCGAAATATAGGCTTATTTTGTGTTAGCTTCAGAGAACCCAG 44564
 OY 927 ACCTTTCCAAAGCTCAGGATTCCTTGAAAAGTTGAAAATTTGATGACTTCAAGCTGA 986
 Db 44563 ACCTTTCCAAAGCTCAGGATTCCTTGAAAAGTTGAAAATTTGATGACTTCAAGCTGA 44504
 OY 987 AGACTTTCAGATTTGAAGGTTACATTCGCAATCCACTATTTAAATATGGAATGGCTGTTA 1046
 Db 44503 AGACTTTCAGATTTGAAGGTTACATTCGCAATCCACTATTTAAATATGGAATGGCTGTTA 44444
 OY 1047 GGGTGTCTTCAAGAGAGCTGGAAGGATTTGTCAGCTTTAGGGGTGGGATGCG 1106
 Db 44443 GGGTGTCTTCAAGAGAGCTGGAAGGATTTGTCAGCTTTAGGGGTGGGATGCG 44384
 OY 1107 AGGTAAAGTCTTTTGTCTTAAAGAAANNAAGAACTAGTCAAAAATCTGTCGGTAC 1166
 Db 44383 AGGTAAAGTCTTTTGTCTTAAAGAAANNAAGAACTAGTCAAAAATCTGTCGGTAC 44324
 OY 1167 CTATCACTTATTTATTTTAAGAGTTCGCACTGCAAAATGTAATCTGTCGAGTTCTT 1226
 Db 44323 CTATCACTTATTTATTTTAAGAGTTCGCACTGCAAAATGTAATCTGTCGAGTTCTT 44264
 OY 1227 CCATATTAAGAGCTTTGATTAACCTGAGAGGATCTGCAATGCTGAGTTATGA 1286
 Db 44263 CCATATTAAGAGCTTTGATTAACCTGAGAGGATCTGCAATGCTGAGTTATGA 44204
 OY 1287 CAAAGTGAAGATGATGATGATGCTCTTAGCAAAAACATGATATGTCATTTCAATC 1346
 Db 44203 CAAAGTGAAGATGATGATGATGCTCTTAGCAAAAACATGATATGTCATTTCAATC 44144
 OY 1347 CCACGTACTTATTAAGAGAGTGGTGAATTTCAAGACTATTTTGAATATTTTGA 1406
 Db 44143 CCACGTACTTATTAAGAGAGTGGTGAATTTCAAGACTATTTTGAATATTTTGA 44084
 OY 1407 TATTTTAAGATTTTCAACAGCTATTTCCCTCAATTCGAGGAGCTGAGTAACCATCA 1466
 Db 44083 TATTTTAAGATTTTCAACAGCTATTTCCCTCAATTCGAGGAGCTGAGTAACCATCA 44024
 OY 1467 TCATGATGAGAGTGGTGAATGAACTTTANAGTTTATATGCTATATTAAGA 1526
 Db 44023 TCATGATGAGAGTGGTGAATGAACTTTATAGCTTTATATGCTATATTAAGA 43964
 OY 1527 AGTGTCTGC 1536
 Db 43963 AGTGTCTGC 43954

RESULT 5
 ABA93402/c
 ID ABA93402 standard; DNA; 45989 BP.
 XX ABA93402;
 AC ABA93402;
 XX ABA93402;
 DT 22-APR-2002 (first entry)
 XX 22-APR-2002 (first entry)
 DE Human rTS-beta gene SEQ ID NO:6.
 XX Human rTS-beta gene SEQ ID NO:6.
 KW Human; rTS-alpha; rTS-beta; rTS; thymidylate synthase; chromosome 18;
 KW mutant thymidylate synthase; enzyme; antitumor; antidepressant;
 KW neuroleptic; neurotropic; tranquilizer; gene therapy; hypomania;
 KW neuropsychiatric disorder; bipolar affective disorder; schizophrenia;
 KW severe bipolar affective disorder; bipolar affective disorder;
 KW major depression; attention deficit disorder; schizoaffective disorder;
 KW gene; ds.
 XX Homo sapiens.
 OS Homo sapiens.
 XX US632344-B1.
 PN US632344-B1.
 XX 27-NOV-2001.
 PD 27-NOV-2001.
 XX 05-NOV-1997; 97US-0965048.
 PF 05-NOV-1997; 97US-0965048.
 XX 05-NOV-1997; 97US-0965048.
 PR 05-NOV-1997; 97US-0965048.
 XX 05-NOV-1997; 97US-0965048.
 PA (MILL-) MILLENNIUM PHARM INC.
 PA (REGC) UNIT CALIFORNIA.
 XX Chen H, Freimer NB;
 PI Chen H, Freimer NB;
 XX WPI; 2002-112960/15.
 DR WPI; 2002-112960/15.
 XX Treating an rTS (mutant form of thymidylate synthase) - mediated
 PT neuropsychiatric disorder, e.g., bipolar affective disorder, comprises
 PT administering a compound, identified by an rTS assay, that interferes
 PT with rTS gene product -
 PT with rTS gene product -
 XX Disclosure; Fig 4A-S; 110pp; English.
 PS The present invention describes a method for treating an rTS (a mutant
 CC form of thymidylate synthase) mediated neuropsychiatric disorder. The
 CC method comprises identifying a compound (C) that can be used to treat
 CC the disorder, by contacting a small molecule test compound (I) with an
 CC rTS protein (II) comprising a sequence of either 361 or 416 amino acids
 CC (AB05587 or AB05588) determining whether (I) binds to (II). (C) has
 CC antitumor, antidepressant, neuroleptic, neurotropic and tranquilizer
 CC activities. rTS polynucleotide sequences can be used in gene therapy.
 CC The method is useful in identifying compounds for treating rTS-mediated
 CC neuropsychiatric disorders including bipolar affective disorder, e.g.,
 CC severe bipolar affective (mood) disorder, bipolar affective (mood)
 CC disorder with hypomania, major depression, schizophrenia, attention
 CC deficit disorder and schizoaffective disorder. The present sequence
 CC represents human rTS-beta gene which is used in the exemplification of
 CC the present invention. rTS has been located to chromosome 18.
 XX Sequence 45989 BP; 12183 A; 10457 C; 10488 G; 12799 T; 62 other;
 SQ Sequence 45989 BP; 12183 A; 10457 C; 10488 G; 12799 T; 62 other;
 Query Match 40.9%; Score 627.8; DB 24; Length 45989;
 Best Local Similarity 91.1%; Pred. No. 2,9e-153;
 Matches 665; Conservative 0; Mismatches 65; Indels 0; Gaps 0;
 QY 807 GTACATGATTGGCAGATCAGCGGCTGAAGCCAGGATTTATACACTTTGGGAGA 866
 DB 44956 GAACTTTTGATGATCACTCTGTGTTCTTTTATGATGATGAGGATTTACACAG 44957
 QY 867 TGACATATTTTACCTGATCAGATCAGCGCACTGAAATTTCACTTCAAGGAGAACCCAG 926
 DB 44896 GTCTGACATTTATGGGAAATATAGCCTTATTTTGTTTTACCTTCAAGGAGAACCCAG 44897
 QY 927 ACCTTTCCAAAGCTGAGATTTCTTGAAAAGTTGAGAAAATGATGACTTCAAGACTGA 986

DB 44836 ACCTTTCCAAAGCTGAGATTTCTTGAAAAGTTGAGAAAATGATGACTTCAAGACTGA 44777
 QY 987 AGACTTTGACATTTGAGGATCAGATCAGATCAGATTTTAAATGGAATGGCTGTTA 1046
 DB 44776 AGACTTTGACATTTGAGGATCAGATCAGATCAGATTTTAAATGGAATGGCTGTTA 44717
 QY 1047 GGGTGTCTTTGAAAGGAGCTGAAAGATTTGTGAGTCTTTAGGGGTGGGCTGAGATCCG 1106
 DB 44716 GGGTGTCTTTGAAAGGAGCTGAAAGATTTGTGAGTCTTTAGGGGTGGGCTGAGATCCG 44657
 QY 1107 AGGTAAAGTCTTTTGTCTGAAAGAAAGGAATAGTCAAAAATGTCGCTGAC 1166
 DB 44656 AGGTAAAGTCTTTTGTCTGAAAGAAAGGAATAGTCAAAAATGTCGCTGAC 44597
 QY 1167 CATCAGTTTATTTTAAAGATGTTGCACTGCAATGAAATGCTGCAATGCTTT 1226
 DB 44596 CATCAGTTTATTTTAAAGATGTTGCACTGCAATGAAATGCTGCAATGCTTT 44537
 QY 1227 CCATATTAAGGCTTTGATTAATCTGAGGATCTGACATGCTGAGTTATGA 1286
 DB 44536 CCATATTAAGGCTTTGATTAATCTGAGGATCTGACATGCTGAGTTATGA 44477
 QY 1287 CAAAGTGAAGAAATGAAATGATGCTCTTTAGCAAAAATGATGCTTTGATC 1346
 DB 44476 CAAAGTGAAGAAATGAAATGATGCTCTTTAGCAAAAATGATGCTTTGATC 44417
 QY 1347 CACGCTACTTAAAGAAAGTGGATTTCAACAATTTTGGAAATTTTGGAA 1406
 DB 44416 CACGCTACTTAAAGAAAGTGGATTTCAACAATTTTGGAAATTTTGGAA 44357
 QY 1407 TATTTTAAAGTTTCAAGCTATTCCTCAATCTGAGGAGCTGAGTACACATGA 1466
 DB 44356 TATTTTAAAGTTTCAAGCTATTCCTCAATCTGAGGAGCTGAGTACACATGA 44297
 QY 1467 TCATGATGAGAGTGGTTATGAATTTANAGTTGTTTATGTTGCTATTAATGA 1526
 DB 44296 TCATGATGAGAGTGGTTATGAATTTANAGTTGTTTATGTTGCTATTAATGA 44237
 QY 1527 AGTGTCTGCG 1536
 DB 44236 AGTGTCTGCG 44227

RESULT 6
 AAF31109
 ID AAF31109 standard; cDNA; 18596 BP.
 XX AAF31109;
 AC AAF31109;
 XX AAF31109;
 DT 27-APR-2001 (first entry)
 DE Thymidylate synthase coding sequence.
 KW Analyte-binding enzyme; analyte analysis; ss.
 XX Homo sapiens.
 OS Homo sapiens.
 XX WO200102600-A2.
 PN WO200102600-A2.
 XX 11-JAN-2001.
 PD 11-JAN-2001.
 XX 30-JUN-2000; 2000WO-US18057.
 PF 30-JUN-2000; 2000WO-US18057.
 XX 06-JUL-1999; 99US-0347878.
 PR 06-DEC-1999; 99US-0457205.
 XX (GEAT) GEN ATOMICS.
 PA (GEAT) GEN ATOMICS.
 PI Yuan C;
 XX Yuan C;
 DR WPI; 2001-071583/08.

PT Assaying method, useful for prognosis and diagnosis of disease,
PT comprises contacting sample with a mutant analyte-binding enzyme and
PT detecting binding -

XX Disclosure, Page -; 187pp; English.

XX The present invention relates to a method for assaying an analyte in a
CC sample comprising: contacting the sample with a mutant analyte-binding
CC enzyme which has binding affinity for the analyte or an immediate
CC analyte enzymatic conversion product but has attenuated catalytic
CC activity; and detecting resulting binding. The method is useful in
CC monitoring biological systems/processes, or prognosis/diagnosis of
CC disease caused by imbalances of the analytes. The present sequence is
CC a coding sequence used in the present invention.
CC Note: the present sequence is not shown in the specification, but was
CC from Genbank, using information given in the specification.

XX Sequence 18596 BP; 4521 A; 3991 C; 4479 G; 5605 T; 0 other;

Query Match 39.8%; Score 611.8; DB 22; Length 18596;
Best Local Similarity 90.4%; Pred. No. 2.9e-149;
Matches 665; Conservative 0; Mismatches 65; Indels 6; Gaps 1;

QY 807 GTACATGATTCGCGACATCAAGGCGCTGAAGCCAGTACTTTATACACTTTGGGAGA 866

DB 15510 GAACCTTGTTGATCACHCTCTGTGACTTTGTTTACAGGACATGAGGAGATTTCAACAG 15569

QY 867 TGCACATATTTTACCTGATCACAATGAGCCACTGAATAATTCAGCTTCAGGAGAACCCAG 926

DB 15570 GTGCGACATATATGCGCAATATATGCGCTTATTTTGTGTTTACCTTCAGCGAGAACCCAG 15629

QY 927 AACCTTCCGAAGTCAGATCTCTCGAAAGTGAAGAAATGATGATCTTCAAGCTGA 986

DB 15630 ACCCTTCCGAAGTCAGATCTCTCGAAAGTGAAGAAATGATGATCTTCAAGCTGA 15689

QY 987 AGACTTCGATTTGAAGGGGACATCCGACATCCACTATTTAAATGCAATGCGTGTGTTA 1046

DB 15690 AGACTTCGATTTGAAGGGGACATCCGACATCCACTATTTAAATGCAATGCGTGTGTTA 15749

QY 1047 GGGTGCCTTTCAAGAGAGCTNGAAGGATTTGCTAGTCTTTAGGGGTTGGGCTGAGTCCG 1106

DB 15750 GGGTGCCTTTCAAGAGAGCTNGAAGGATTTGCTAGTCTTTAGGGGTTGGGCTGAGTCCG 15809

QY 1107 AGGTAAAGTCTTTTGTCTTAAAGAAAGAAAGAACTAGGTGAAATCTGTCGCTGAC 1166

DB 15810 AGGTAAAGTCTTTTGTCTTAAAGAAAGAAAGAACTAGGTGAAATCTGTCGCTGAC 15869

QY 1167 CATAGATTTAATTTTAAAGAGTGTGCACTGGGAATGTAACTGTGCGAGTCTTT 1226

DB 15870 CATAGATTTAATTTTAAAGAGTGTGCACTGGGAATGTAACTGTGCGAGTCTTT 15929

QY 1227 CCATATATTAAGGCTTTAGGTTAACTCACTGAGGGGTATCTGACATGCTGAGTTATGAA 1286

DB 15930 CCATATATTAAGGCTTTAGGTTAACTCACTGAGGGGTATCTGACATGCTGAGTTATGAA 15989

QY 1287 CAAAGTGAAGAAATGAATGTATGTCTTTAGCAAAAACATGATGTGCAATTTCAATC 1346

DB 15990 CAAAGTGAAGAAATGAATGTATGTCTTTAGCAAAAACATGATGTGCAATTTCAATC 16049

QY 1347 CCACTACTTTTAAAGAAAGTGTGTAATTTCAACACTTTTGTGAAATTTTGTGAA 1406

DB 16050 CCACTACTTTTAAAGAAAGTGTGTAATTTCAACACTTTTGTGAAATTTTGTGAA 16109

QY 1407 TATTTTAAAGATTTCAACAGCTATCCCTCAAACTGAAGGAGCTGAGTAAACACACGA 1466

DB 16110 TATTTTAAAGATTTCAACAGCTATCCCTCAAACTGAAGGAGCTGAGTAAACACACGA 16169

QY 1467 TCATGATGTAGAGTGTGTTATGAATCTTTTAAAGTGTGTTATATGTGCTATAA 1520

DB 16170 TCATGATGTAGAGTGTGTTATGAATCTTTTAAAGTGTGTTATATGTGCTATAA 16229

QY 1521 TAAAGAGTGTCTGCG 1536

DB 16230 TAAAGAGTGTCTGCG 16245

RESULT 7
AAC91215
ID AAC91215 standard; DNA; 18596 BP.

XX AAC91215;

DT 20-MAR-2001 (first entry)

XX Human thymidylate synthase gene SEQ ID NO: 11.

XX Human; schizophrenia; developmental disorder; spina bifida cystica;

XX Tourette's syndrome; bipolar illness; autism; conduct disorder;

XX attention deficit disorder; obsessive compulsive disorder;

XX chronic multiple tic syndrome; learning disorder; polymorphism; ds.

XX Homo sapiens.

XX WO200071754-A1.

XX 30-NOV-2000.

XX 24-MAY-2000; 2000WO-US14354.

XX 25-MAY-1999; 99US-0318448.

XX (UYNE-) UNIV NEW JERSEY MEDICINE & DENTISTRY.

XX Johnson WG, Stenroos ES;

XX WPI; 2001-025174/03.

XX Diagnosing a developmental disorder, e.g. schizophrenia, by forming

XX datasets (DS) of genetic (e.g. genotypes of folate metabolism alleles)

XX and environmental variables affecting an individual and then comparing

XX these DS with reference DS -

XX

XX Disclosure, Page 125-131; 156pp; English.

XX The present invention provides a novel method of estimating the

XX susceptibility of an individual to a developmental disorder using genetic

XX and environmental variables. The method can be used in the diagnosis,

XX prevention and treatment of disorders such as schizophrenia, spina bifida

XX cystica, Tourette's syndrome, bipolar illness, autism, conduct disorders,

XX attention deficit disorder, obsessive compulsive disorder, chronic

XX multiple tic syndrome and learning disorders such as dyslexia.

XX Sequence 18596 BP; 4521 A; 3991 C; 4479 G; 5605 T; 0 other;

XX

XX Query Match 39.8%; Score 611.8; DB 22; Length 18596;

XX Best Local Similarity 90.4%; Pred. No. 2.9e-149;

XX Matches 665; Conservative 0; Mismatches 65; Indels 6; Gaps 1;

QY 807 GTACATGATTCGCGACATCAAGGCGCTGAAGCCAGTACTTTATACACTTTGGGAGA 866

DB 15510 GAACCTTGTTGATCACHCTCTGTGACTTTGTTTACAGGACATGAGGAGATTTCAACAG 15569

QY 867 TGCACATATTTTACCTGATCACAATGAGCCACTGAATAATTCAGCTTCAGGAGAACCCAG 926

DB 15570 GTGCGACATATATGCGCAATATATGCGCTTATTTTGTGTTTACCTTCAGCGAGAACCCAG 15629

QY 927 AACCTTCCGAAGTCAGATCTCTCGAAAGTGAAGAAATGATGATCTTCAAGCTGA 986

DB 15630 ACCCTTCCGAAGTCAGATCTCTCGAAAGTGAAGAAATGATGATCTTCAAGCTGA 15689

QY 987 AGACTTCGATTTGAAGGGGACATCCGACATCCACTATTTAAATGCAATGCGTGTGTTA 1046

DB 15690 AGACTTCGATTTGAAGGGGACATCCGACATCCACTATTTAAATGCAATGCGTGTGTTA 15749

QY 1047 GGGTGCCTTTCAAGAGAGCTNGAAGGATTTGCTAGTCTTTAGGGGTTGGGCTGAGTCCG 1106

Db 15750 GGGTGGCTTTCAAGAGGAGCTGCAAGATATTTGTCACTTTAGGGGTTGGGCTGATGCGG 15809
Qy 1107 AGGTAAAGTCTTTTGGCTTAAAGAAAGAACTAGATCAAAATCTGCTCCGTGAC 1166
Db 15810 AGGTAAAGTCTTTTGGCTTAAAGAAAGAACTAGATCAAAATCTGCTCCGTGAC 15869
Qy 1167 CTATCAGTTATTAATTTTAAAGATGTTGCACTGGCAAAATGTAATCTGCTGCTTT 1226
Db 15870 CTATCAGTTATTAATTTTAAAGATGTTGCACTGGCAAAATGTAATCTGCTGCTTT 15929
Qy 1227 CCAATATTAAGAGCTTTGAGTAACTCACTGAGGGTATCTGCAAAATGCTGAGGTTATGAA 1286
Db 15930 CCAATATTAAGAGCTTTGAGTAACTCACTGAGGGTATCTGCAAAATGCTGAGGTTATGAA 15989
Qy 1287 CAAAGTGAAGAGATGAATATATGCTCTTACCAAAACATGTAATGCAATTTCAATC 1346
Db 15990 CAAAGTGAAGAGATGAATATATGCTCTTACCAAAACATGTAATGCAATTTCAATC 16049
Qy 1347 CCAAGTGAAGAGATGAATATATGCTCTTACCAAAACATGTAATGCAATTTCAATC 1406
Db 16050 CCAAGTGAAGAGATGAATATATGCTCTTACCAAAACATGTAATGCAATTTCAATC 16109
Qy 1407 TATTTAAGATTTCAAGAGTATTCCTCAAAATCTGAGGAGCTGATTAACCATGGA 1466
Db 16110 TATTTAAGATTTCAAGAGTATTCCTCAAAATCTGAGGAGCTGATTAACCATGGA 16169
Qy 1467 TCATGATGTAAGAGTGGTATGAACT-----TTAATGTTGTTTAAATGTTGCTATTA 1520
Db 16170 TCATGATGTAAGAGTGGTATGAACTTTAAAGTTATATGTTTAAATGTTGCTATTA 16229
Qy 1521 TAAAGAGTGTCTGC 1536
Db 16230 TAAAGAGTGTCTGC 16245

RESULT 8
ABN95092
ID ABN95092 standard; DNA; 18596 BP.
XX
AC ABN95092;
XX
XX 13-AUG-2002 (first entry)
XX
XX
XX Gene #1590 used to diagnose liver cancer.
XX
XX Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
XX
XX metastatic liver tumor; cytostatic; expression profile; disease state;
XX
XX disease progression; drug toxicity; drug efficacy; drug metabolism.
XX
XX Homo sapiens.
XX
XX
XX WO200229103-A2.
XX
XX
XX 11-APR-2002.
XX
XX
XX 02-OCT-2001; 2001MO-US30589.
XX
XX
XX 02-OCT-2000; 2000US-237054P.
XX
XX
XX (GENE-) GENE LOGIC INC.
XX
XX
XX Horne D, Alvares C, Petes-Da-Silva S, Vockley JG;
XX
XX
XX WPI; 2002-426119/45.
XX
XX
XX
XX Diagnosing and detecting the progression of liver cancer.
XX
XX hepatocellular carcinoma or metastatic liver tumor in a patient,
XX
XX involves detecting the level of expression of two or more genes in a
XX
XX liver tissue sample -
XX
XX
XX Claim 1; SEQ ID NO 1590; 298bp; English.
XX
XX
XX The invention relates to a novel method for diagnosing and detecting the

CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
CC tumor in a patient, and differentiating metastatic liver cancer from
CC hepatocellular carcinoma in a patient, involving detecting the level of
CC expression of two or more genes represented in ABN95092-ABN9745 in a
CC tissue sample. The method of the invention has hepatotropic, and
CC cytostatic activity. The method is useful for diagnosing and detecting
CC the progression of liver cancer, hepatocellular carcinoma and metastatic
CC liver carcinoma in a patient. The method is useful for identifying
CC expression profiles which serve as useful diagnostic markers as well as
CC markers that can be used to monitor disease states, disease progression,
CC drug toxicity, drug efficacy and drug metabolism.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pat_sequences.
XX

Sequence 18596 BP; 4521 A; 3991 C; 4479 G; 5605 T; 0 other;

Query Match 39.8%; Score 611.8; DB 24; Length 18596;

Best Local Similarity 90.4%; Pred.No.2,9e-149;

Matches 665; Conservative 0; Mismatches 65; Indels 6; Gaps 1;

Qy 807 GTACATGATTCGCGACATCAAGGCGCTGAGCCAGGATCTTATACACATTTGGGAGA 866
Db 15510 GAACCTTGTATCAATCACTCTGTGATCTTTTACAGAGATGAGACATTAACACAG 15569
Qy 867 TCGACATTAATTAAGTATCAATCACTGAGCCAGTGAATTAATCACTTCAAGAGAACCCAG 926
Db 15570 GTGCGATTAATTAAGTATCAATCACTGAGCCAGTGAATTAATCACTTCAAGAGAACCCAG 15629
Qy 927 ACCCTTCCCAAGAGTCAAGATTCCTGAAAAGTTGAGAAAATGATGACTTCAAAAGCTGA 986
Db 15630 ACCCTTCCCAAGAGTCAAGATTCCTGAAAAGTTGAGAAAATGATGACTTCAAAAGCTGA 15689
Qy 987 AGACTTTCAGATGAGGATCAATCCGATCACAATTAAATGGAATGGCTGTTTA 1046
Db 15690 AGACTTTCAGATGAGGATCAATCCGATCACAATTAAATGGAATGGCTGTTTA 15749
Qy 1047 GGGTCTTTCAAGAGCTGAGGATGATGCTCTTTAGGGTGGGCTGATGTCGG 1106
Db 15750 GGGTCTTTCAAGAGCTGAGGATGATGCTCTTTAGGGTGGGCTGATGTCGG 15809
Qy 1107 AGGTAAAGTCTTTTGGCTTAAAGAAAGAACTAGATCAAAATCTGCTCCGTGAC 1166
Db 15810 AGGTAAAGTCTTTTGGCTTAAAGAAAGAACTAGATCAAAATCTGCTCCGTGAC 15869
Qy 1167 CTATCAGTTATTAATTTTAAAGATGTTGCACTGGCAAAATGTAATCTGCTGCTTT 1226
Db 15870 CTATCAGTTATTAATTTTAAAGATGTTGCACTGGCAAAATGTAATCTGCTGCTTT 15929
Qy 1227 CCAATATTAAGAGCTTTGAGTAACTCACTGAGGGTATCTGCAAAATGCTGAGGTTATGAA 1286
Db 15930 CCAATATTAAGAGCTTTGAGTAACTCACTGAGGGTATCTGCAAAATGCTGAGGTTATGAA 15989
Qy 1287 CAAAGTGAAGAGATGAATATATGCTCTTACCAAAACATGTAATGCAATTTCAATC 1346
Db 16050 CAAAGTGAAGAGATGAATATATGCTCTTACCAAAACATGTAATGCAATTTCAATC 16049
Qy 1347 CCAAGTGAAGAGATGAATATATGCTCTTACCAAAACATGTAATGCAATTTCAATC 1406
Db 16110 CCAAGTGAAGAGATGAATATATGCTCTTACCAAAACATGTAATGCAATTTCAATC 16109
Qy 1407 TATTTAAGATTTCAAGAGTATTCCTCAAAATCTGAGGAGCTGATTAACCATGGA 1466
Db 16110 TATTTAAGATTTCAAGAGTATTCCTCAAAATCTGAGGAGCTGATTAACCATGGA 16169
Qy 1467 TCATGATGTAAGAGTGGTATGAACT-----TTAATGTTGTTTAAATGTTGCTATTA 1520
Db 16170 TCATGATGTAAGAGTGGTATGAACTTTAAAGTTATATGTTTAAATGTTGCTATTA 16229
Qy 1521 TAAAGAGTGTCTGC 1536
Db 16230 TAAAGAGTGTCTGC 16245

	QY	1321	CAAAACATGTATGTCATTTCATCCACAGTCTTAATAAAGAGTTGGTAATTCAC	1380
	Dd	1321	CAAAACATGTATGTCATTTCATCCACAGTCTTAATAAAGAGTTGGTAATTCAC	1380
	QY	1381	AAGCTATTTTGGAAATATTTTAGAATTTTTAGAATTTCCAAAGCTATTCCTCAAAAT	1440
	Dd	1381	AGCTATTTTGGAAATATTTTAGAATTTTTAGAATTTCCAAAGCTATTCCTCAAAAT	1440
	QY	1441	CTGAGGAGAGCTGACTAACACCATCATCATGATGTAGATGAGTGATGAAGCTTGA	1496
	Dd	1441	CTGAGGAGAGCTGACTAACACCATCATCATGATGTAGATGAGTGATGAAGCTTGA	1496
		RESULT 12		
		US-60-412-049-350		
		; Sequence 350, Application US/60412049		
		; GENERAL INFORMATION:		
		; APPLICANT: Baker, Joffre B.		
		; APPLICANT: Cronin, Maureen T.		
		; APPLICANT: Kiefer, Michael C.		
		; APPLICANT: Shak, Steve		
		; APPLICANT: Walker, Michael Graham		
		; TITLE OF INVENTION: GENE EXPRESSION PROFILING IN BIOPSIED TUMOR TISSUES		
		; FILE REFERENCE: GENHR.004PR		
		; CURRENT APPLICATION NUMBER: US/60/412,049		
		; CURRENT FILING DATE: 2002-09-18		
		; NUMBER OF SEQ ID NOS: 384		
		; SOFTWARE: FastSeq for Windows Version 4.0		
		; SEQ ID NO 350		
		; LENGTH: 1536		
		; TYPE: DNA		
		; ORGANISM: Homo sapiens		
		US-60-412-049-350		
		Query Match	90.8%; Score 1394; DB 85; Length 1536;	
		Best Local Similarity 99.9%; Pct. No. 0;		
		Matches 1494; Conservative 0; Mismatches 2; Indels 0; Gaps		
	QY	1	GGGGGGGGGGGGGACCACTTGGCTCGCTCCGTCGGCGGCGGCACCTTGGCTCGCTCGGT	60
	Dd	1	GGGGGGGGGGGGGACCACTTGGCTCGCTCCGTCGGCGGCGGCACCTTGGCTCGCTCGGT	60
	QY	61	CCGCGCGCGCACCTTCCGCTGCCCGCGGTCCCGCGCGCGCGCGCCTCATGCTTGCGCGGC	120
	Dd	61	CCGCGCGCGCACCTTCCGCTGCCCGCGGTCCCGCGCGCGCGCGCCTCATGCTTGCGCGGC	120
	QY	121	TGGAGAGTGGCGCGCGCGCTTACCCTCCCGCGCGCACAGAGCGGGACAACCGCGCGGT	180
	Dd	121	TGGAGAGTGGCGCGCGCGCTTACCCTCCCGCGCGCACAGAGCGGGACAACCGCGCGGT	180
	QY	181	CCGCGCGCACGGGAGACTGACAGTACCTGCGGAGTCAGATCCAAACATCTCTCCGCTGCGCGT	240
	Dd	181	CCGCGCGCACGGGAGACTGACAGTACCTGCGGAGTCAGATCCAAACATCTCTCCGCTGCGCGT	240
	QY	241	AGGAAGACGACCGCACCGGACCGGACCGCTGTGTGATTCGGCATGACAGGCGCGCTAC	300
	Dd	241	AGGAAGACGACCGCACCGGACCGGACCGCTGTGTGATTCGGCATGACAGGCGCGCTAC	300
	QY	301	AGCTTGAGAGATGAATTCCTCTCTCTCTGACACCAACAGCTGTGTTCTGGAAAGGATTTTG	360
	Dd	301	AGCTTGAGAGATGAATTCCTCTCTCTCTGACACCAACAGCTGTGTTCTGGAAAGGATTTTG	360
	QY	361	GAGAGATTGCTGTGTTTATTAAGGATCCACAAANGCTAAAGAGCTGTTCACAGGGA	420
	Dd	361	GAGAGATTGCTGTGTTTATTAAGGATCCACAAANGCTAAAGAGCTGTTCACAGGGA	420
	QY	421	GTGAAATCTGGGATGCCAAGGATTCACGACTTTTGGACAACCTCGGATTTTCAC	480
	Dd	421	GTGAAATCTGGGATGCCAAGGATTCACGACTTTTGGACAACCTCGGATTTTCAC	480
	QY	481	AGAGAAAGAGGAGACTTGGGCGCCAGATTATGAGCTTCAGTGGAGAGCAATTTTGGGCGAGA	540
	Dd	481	AGAGAAAGAGGAGACTTGGGCGCCAGATTATGAGCTTCAGTGGAGAGCAATTTTGGGCGAGA	540

Db	693	AAGAGAAAGAGGACCTTGGGCGCCAGTTTATGCTCCAGTGGAGAGCATTTTGGGAGCAGA	752
Qy	541	TACAGAGATATGGAATCCAGATTATTTCAGACAGGAGTTTACCAATCTCAAGAGATGAT	600
Db	753	TTCAGAGATATGGAATCAGATTATTTCAGACAGGAGGATTTGACCAATCTCAAGAGATGAT	812
Qy	601	GACACATTCAAAACCAACCTGACAGCAGAAATCATCATGTGGGCTTGGAAATCCAAACA	660
Db	813	GACACATTCAAAACCAACCTGACAGCAGAAATCATCATGTGGGCTTGGAAATCCAAACA	872
Qy	661	GATCTTCCTGTATGGGCGCTTCCCATCATGCGCTCTGCACAGTTTATGTGGTAAAC	720
Db	873	GATCTTCCTGTATGGGCGCTTCCCATCATGCGCTCTGCACAGTTTATGTGGTAAAC	932
Qy	721	AATGAGCTGTCTCCAGAGCTGTACAGAAATGAGGAGCAATGGGCGCTGTGTGCTCTTC	780
Db	933	AATGAGCTGTCTCCAGAGCTGTACAGAAATGAGGAGCAATGGGCGCTGTGTGCTCTTC	992
Qy	781	AACATCCGACGTACGCGCTGTCTCATGTAAGTGGGACATCCAGGCTCTGAAAGCCA	840
Db	993	AACATCCGACGTACGCGCTGTCTCATGTAAGTGGGACATCCAGGCTCTGAAAGCCA	1052
Qy	841	GGTGCATTATACACACTTTTGGAGATGTACATAATTAACCTGATCATCTGAGCCACTG	900
Db	1053	GGTGCATTATACACACTTTTGGAGATGTACATAATTAACCTGATCATCTGAGCCACTG	1112
Qy	901	AAATATGAGCTTACGCGGAATCCGACACTTCCCAAGCTCAGAGATTTCTGAAAAGTT	960
Db	1113	AAATATGAGCTTACGCGGAATCCGACACTTCCCAAGCTCAGAGATTTCTGAAAAGTT	1172
Qy	961	GAGAAATGTATGATCATCTCAAAAGCTGGAACCTTCAGATTGAAGGATCAATCCGACATCA	1020
Db	1173	GAGAAATGTATGATCATCTCAAAAGCTGGAACCTTCAGATTGAAGGATCAATCCGACATCA	1232
Qy	1021	ACTATTAAATGGAATGGCTGTGTAGGGTCTTTCAAAGAGCTTGAAGAGATATGTCA	1080
Db	1233	ACTATTAAATGGAATGGCTGTGTAGGGTCTTTCAAAGAGCTTGAAGAGATATGTCA	1292
Qy	1081	GTCTTTAGGGGTTGGGCTGTGATGTCCGAGGTAAGTCTTTTGTCTTAAAGAAAGAG	1140
Db	1293	GTCTTTAGGGGTTGGGCTGTGATGTCCGAGGTAAGTCTTTTGTCTTAAAGAAAGAG	1352
Qy	1141	AATCTAGGTCAAAAATGTGTCGCTGACCTTACAGTATTAATTTTAAGAGATGTGCACT	1200
Db	1353	AATCTAGGTCAAAAATGTGTCGCTGACCTTACAGTATTAATTTTAAGAGATGTGCACT	1412
Qy	1201	GGCAATGTAACTGTGCCAGTTCTTTTCCATTAATTAAGGCTTTGAGTTAACTCATGTAGG	1260
Db	1413	GGCAATGTAACTGTGCCAGTTCTTTTCCATTAATTAAGGCTTTGAGTTAACTCATGTAGG	1472
Qy	1261	GTATCTGACATCTGTAGGTTATATGAACAAAGTAGAGAAATGAATGTGTGCTGTGAG	1320
Db	1473	GTATCTGACATCTGTAGGTTATATGAACAAAGTAGAGAAATGAATGTGTGCTGTGAG	1532
Qy	1321	CAAAAACATGATATGTCAATTCATCAATCCAGATTAATTAAGAAAGTTGTGTGAATTCAC	1380
Db	1533	CAAAAACATGATATGTCAATTCATCAATCCAGATTAATTAAGAAAGTTGTGTGAATTCAC	1592
Qy	1381	AAGCTATTTTGGAAATATTTTGAATATTTTAAAGATTTTCAAGACTATTTCTCCAAAT	1440
Db	1593	AAGCTATTTTGGAAATATTTTGAATATTTTAAAGATTTTCAAGACTATTTCTCCAAAT	1652
Qy	1441	CTGAGAGAGCTGTAGTACACATGTGATCATGATGTATGAAGTGTGTATGAACCTTA	1496
Db	1653	CTGAGAGAGCTGTAGTACACATGTGATCATGATGTATGAAGTGTGTATGAACCTTA	1708

QY 853 CACACTTGGGAGATGCACATATTTACCTGAATCACTAGACGCACCTGAATAATTGAGCTT 912
DB 887 CACACTTGGGAGATGCACATATTTACCTGAATCACTAGACGCACCTGAATAATTGAGCTT 946
QY 913 CAGGAGAACCCAGACTTTTCCCAAGCTCAGGATCTTCCAAAGTGAAGAAATGAT 972
DB 947 CAGGAGAACCCAGACTTTTCCCAAGCTCAGGATCTTCCAAAGTGAAGAAATGAT 1006
QY 973 GACTCAAAAGCTGAAGCTTCCAGATGAAGGGTCAATCCGACCTCAATATTAATG 1032
DB 1007 GACTCAAAAGCTGAAGCTTCCAGATGAAGGGTCAATCCGACCTCAATATTAATG 1066
QY 1033 GAATGAGCTGTTAAGGCTGCTTCCAAAGGCTGGAAGATATTCAGCTTTAGGGT 1092
DB 1067 GAATGAGCTGTTAAGGCTGCTTCCAAAGGCTGGAAGATATTCAGCTTTAGGGT 1126
QY 1093 TGGGCTGAGATGCCGAGTAAAGTCTTTTGTCTTAAAGAAAGAAAGAACTGTGCAAA 1152
DB 1127 TGGGCTGAGATGCCGAGTAAAGTCTTTTGTCTTAAAGAAAGAAAGAACTGTGCAAA 1186
QY 1153 AATGTGCGTGAACCTATCACTATTTAATTTTAAGATGTGCACTGGCAATGTAC 1212
DB 1187 AATGTGCGTGAACCTATCACTATTTAATTTTAAGATGTGCACTGGCAATGTAC 1246
QY 1213 TGTGCCAGTCTTTCATATATAAAGGCTTGAATTAATCACTGAGGATCTGCAAT 1272
DB 1247 TGTGCCAGTCTTTCATATATAAAGGCTTGAATTAATCACTGAGGATCTGCAAT 1306
QY 1273 GCTGAGGTTATGACAAAGAGAGAGATGAATGTATGTGCTTGAAGCAAAACATGTA 1332
DB 1307 GCTGAGGTTATGACAAAGAGAGATGAATGTATGTGCTTGAAGCAAAACATGTA 1366
QY 1333 TGTGCATTTCAATCCGAGTACTTATAAAGAGTGTGTGAATTCACAGCTATTTTG 1392
DB 1367 TGTGCATTTCAATCCGAGTACTTATAAAGAGTGTGTGAATTCACAGCTATTTTG 1426
QY 1393 GAATATTTTGAATTTTAAAGATTTCAAGCATTTCCCTCAATTCGAGGAGCTG 1452
DB 1427 GAATATTTTGAATTTTAAAGATTTCAAGCATTTCCCTCAATTCGAGGAGCTG 1486
QY 1453 AGTAACACCATGATGATGATGAGTGTGTTAAGACTTAA 1496
DB 1487 AGTAACACCATGATGATGATGAGTGTGTTAAGACTTAA 1530

RESULT 15
us-09-652-121-6686
; Sequence 6686, Application US/09652121
; GENERAL INFORMATION:
; APPLICANT: Stefano, Peter
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 1600.1188-001
; CURRENT APPLICATION NUMBER: US/09/652.121
; PRIOR FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/151.129
; NUMBER OF SEQ ID NOS: 7615
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6686
; LENGTH: 1829
; TYPE: DNA
; ORGANISM: Homo sapiens
us-09-652-121-6686

Query Match 90.0%; Score 1382; DB 25; Length 1829;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1482; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 13 ACACCTGGGCTGCTCCGCTCCCGCGCGCACCTGAGCTGCTCCGCTCCGCGCGGCA 72
DB 47 ACACCTGGGCTGCTCCGCTCCCGCGCGCACCTGAGCTGCTCCGCTCCGCGCGGCA 106

QY 73 CTTCGCTGCTCCGCTCCCGCGCGCACCTGAGCTGCTCCGCTCCGCGCGGCA 132
DB 107 CTTCGCTGCTCCGCTCCCGCGCGCACCTGAGCTGCTCCGCTCCGCGCGGCA 166
QY 133 CCGCGGCTTGTCCCGCGCGCGCACAGAGCGGAGCGCGAGCGCTCCGCGCGGCA 192
DB 167 CCGCGGCTTGTCCCGCGCGCGCACAGAGCGGAGCGCGAGCGCTCCGCGCGGCA 226
QY 193 GAGCTGAGTACTTGGGGAGATCCAAATCTCTCGCTGCGGCTCAGAGAGAGCAG 252
DB 227 GAGCTGAGTACTTGGGGAGATCCAAATCTCTCGCTGCGGCTCAGAGAGAGCAG 286
QY 253 CCGAGCGGACCGGCAACCTCTGCTGATTCGAGATGCGAGCGGCTGAGGAT 312
DB 287 CCGAGCGGACCGGCAACCTCTGCTGATTCGAGATGCGAGCGGCTGAGGAT 346
QY 313 GAATTCCTCTGCTGAGCAACAAAGTGTCTGGAAGGGTGTGGAAGAGTGTCTG 372
DB 347 GAATTCCTCTGCTGAGCAACAAAGTGTCTGGAAGGGTGTGGAAGAGTGTCTG 406
QY 373 TGGTTATCAAGGATCACAAAAGCTTAAAGCTTCTTCAAGGAGTGAATCTG 432
DB 407 TGGTTATCAAGGATCACAAAAGCTTAAAGCTTCTTCAAGGAGTGAATCTG 466
QY 433 GATGCCAATGATCCGAGACTTTTGAAGAGCTTGGGATTTTCCACAGAGAGAGG 492
DB 467 GATGCCAATGATCCGAGACTTTTGAAGAGCTTGGGATTTTCCACAGAGAGAGG 526
QY 493 GACTTGGGCTCAGTTTATGCTTCCAGTGAAGGCAATTTGGGGCAAAATCAGATG 552
DB 527 GACTTGGGCTCAGTTTATGCTTCCAGTGAAGGCAATTTGGGGCAAAATCAGATG 586
QY 553 GATCAGATTTATGAGGACAGGAGTGAACCACTGCAAAAGAGTGAACCATCAAA 612
DB 587 GATCAGATTTATGAGGACAGGAGTGAACCACTGCAAAAGAGTGAACCATCAAA 646
QY 613 ACCAACCCTGACGACAGAGATCATCATGTGCTTGAATCAAGAGATCTTCTG 672
DB 647 ACCAACCCTGACGACAGAGATCATCATGTGCTTGAATCAAGAGATCTTCTG 706
QY 673 ATGGGCTGCTCCATGACCATGCTTCCAGCTTCAATGTGTGAACAGTACCTGCTC 732
DB 707 ATGGGCTGCTCCATGACCATGCTTCCAGCTTCAATGTGTGAACAGTACCTGCTC 766
QY 733 TGTCACTGTACAGAGATGGGAGCATGGGCTCGGTGCTTCAATGCCAGC 792
DB 767 TGTCACTGTACAGAGATGGGAGCATGGGCTCGGTGCTTCAATGCCAGC 826
QY 793 TAGGCCCTGCTCAGTACATGATTTGGGACATCAACGGGCTGAGGCTGACTTTATA 852
DB 827 TAGGCCCTGCTCAGTACATGATTTGGGACATCAACGGGCTGAGGCTGACTTTATA 886
QY 853 CACACTTGGGAGATGCACATATTTAAGTGAATCACTGAGGCTGAGGCTTCAAGCTT 912
DB 887 CACACTTGGGAGATGCACATATTTAAGTGAATCACTGAGGCTGAGGCTTCAAGCTT 946
QY 913 CAGCGAGAACCCAGACTTTTCCCAAGCTCAGGATCTTCCAAAGTGAAGAAATGAT 972
DB 947 CAGCGAGAACCCAGACTTTTCCCAAGCTCAGGATCTTCCAAAGTGAAGAAATGAT 1006
QY 973 GACTCAAAAGCTGAAGCTTCCAGATGAAGGGTCAATCCGACCTCAATATTAATG 1032
DB 1007 GACTCAAAAGCTGAAGCTTCCAGATGAAGGGTCAATCCGACCTCAATATTAATG 1066
QY 1033 GAATGAGCTGTTAAGGCTGCTTCCAAAGGCTGGAAGATATTCAGCTTTAGGGT 1092
DB 1067 GAATGAGCTGTTAAGGCTGCTTCCAAAGGCTGGAAGATATTCAGCTTTAGGGT 1126
QY 1093 TGGGCTGAGATGCCGAGTAAAGTCTTTTGTCTTAAAGAAAGAAAGAACTGTGCAAA 1152
DB 1127 TGGGCTGAGATGCCGAGTAAAGTCTTTTGTCTTAAAGAAAGAAAGAACTGTGCAAA 1186

QY 443 GATCCGAGACCTTTTGGACAGCCTGGATTCTCAACGAGAAAGAGGAGCTTGG393C 502
DB 241 GATCCGAGACCTTTTGGACAGCCTGGATTCTCAACGAGAAAGAGGAGCTTGG393C 300
QY 503 CAGTTATGAGCTTCCAGTGGAGGAGCTTTGGGCGAGAAATACAGATATGAGATTCAGATT 562
DB 301 CAGTTATGAGCTTCCAGTGGAGGAGCTTTGGGCGAGAAATACAGATATGAGATTCAGATT 360
QY 563 ATTGAGACAGGAGATTGACCACTGCAAAAGATGATTAACACATCAAAACCAACCTTG 622
DB 361 ATTGAGACAGGAGATTGACCACTGCAAAAGATGATTAACACATCAAAACCAACCTTG 420
QY 623 ACACAGAGAAATCATCATGTCCTTGGAAATCAAGAGAAATTCCTCTGATGAGGCGCTG 682
DB 421 ACACAGAGAAATCATCATGTCCTTGGAAATCAAGAGAAATTCCTCTGATGAGGCGCTG 480
QY 683 CTCGATGCAATGCTCTGCGAGTTCTATGAGTGAACAGTGAAGCTGCTG 734
DB 481 CTCGATGCAATGCTCTGCGAGTTCTATGAGTGAACAGTGAAGCTGCTG 532

RESULT 2

US-09-954-531-124
Sequence 124, Application US/09954531
Patent No. US20020165180A1
GENERAL INFORMATION:
APPLICANT: Meaver, Zoe
TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Can
FILE REFERENCE: 689290-77
CURRENT FILING DATE: 2002-05-02
PRIOR FILING DATE: 2002-05-02
PRIOR APPLICATION NUMBER: US/60/233,133
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: US/60/234,009
PRIOR FILING DATE: 2000-09-20
PRIOR APPLICATION NUMBER: US/60/234,034
PRIOR FILING DATE: 2000-09-20
PRIOR APPLICATION NUMBER: US/60/234,509
PRIOR FILING DATE: 2000-09-22
PRIOR APPLICATION NUMBER: US/60/234,567
PRIOR FILING DATE: 2000-09-22
NUMBER OF SEQ ID NOS: 1392
SOFTWARE: Patent version 3.0
SEQ ID NO 124
LENGTH: 18596
TYPE: DNA
ORGANISM: Homo sapiens
US-09-954-531-124

Query Match 31.7%; Score 487; DB 9; Length 18596;

Best Local Similarity 99.7%; Pred. No. 4,5e-249; Indels 0; Gaps 0;

Matches 587; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 908 AGCTTCAGGAGAACCCAGACCTTTCCAAAGCTCAGAGATCTTGGAAATTTGAGAAA 967
DB 15611 AGCTTCAGGAGAACCCAGACCTTTCCAAAGCTCAGAGATCTTGGAAATTTGAGAAA 15670
QY 968 TTGATGACTTCAAGCTGAAGCTTTGAGATGAAGGTACATCCGATCAACTTTA 1027
DB 15671 TTGATGACTTCAAGCTGAAGCTTTGAGATGAAGGTACATCCGATCAACTTTA 15730
QY 1028 AAATGAAATGAGCTGTTAGGAGCTTTCAAGAGCTGAAAGATATTTGCACTTTA 1087
DB 15731 AAATGAAATGAGCTGTTAGGAGCTTTCAAGAGCTGAAAGATATTTGCACTTTA 15790
QY 1088 GGGGTTGGGCTGAGATCCGAGGTAAAGTTCTTTTCTTAAAGAAAGAACTAGG 1147
DB 15791 GGGGTTGGGCTGAGATCCGAGGTAAAGTTCTTTTCTTAAAGAAAGAACTAGG 15850
QY 1148 TCAAAAATCTGTCGATGACATGATTAATTTTAAAGATGTTGCACTGGCAAT 1207

DB 15851 TCAAAAATCTGTCGATGACATGATTAATTTTAAAGATGTTGCACTGGCAAT 15910
QY 1208 GATACGTGCGAGATCTTTCCATTAATAAAGCTTTGATTAATCACTGAGGATCTG 1267
DB 15911 GATACGTGCGAGATCTTTCCATTAATAAAGCTTTGATTAATCACTGAGGATCTG 15970
QY 1268 ACAATGCTGAGGTATGAACAAGTGAAGAAATGAATGATGATGCTCTTGAAGAAAAC 1327
DB 15971 ACAATGCTGAGGTATGAACAAGTGAAGAAATGAATGATGATGCTCTTGAAGAAAAC 16030
QY 1328 ATGTATGCAATTCATCCAGTACTTATTAAGAGGTTGATTAATCAAGACTAT 1387
DB 16031 ATGTATGCAATTCATCCAGTACTTATTAAGAGGTTGATTAATCAAGACTAT 16090
QY 1388 TTTTGAATATTTTGAATATTTTGAATTTTCAAGATTTCCCTCAATCTGAGG 1447
DB 16091 TTTTGAATATTTTGAATATTTTGAATTTTCAAGATTTCCCTCAATCTGAGG 16150
QY 1448 AGCTGAGTAAACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1496
DB 16151 AGCTGAGTAAACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 16199

RESULT 3

US-09-954-531-348
Sequence 348, Application US/09954531
Patent No. US20020165180A1
GENERAL INFORMATION:
APPLICANT: Meaver, Zoe
TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using
FILE REFERENCE: 689290-77
CURRENT FILING DATE: 2002-05-02
PRIOR FILING DATE: 2002-05-02
PRIOR APPLICATION NUMBER: US/60/233,133
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: US/60/234,009
PRIOR FILING DATE: 2000-09-20
PRIOR APPLICATION NUMBER: US/60/234,034
PRIOR FILING DATE: 2000-09-20
PRIOR APPLICATION NUMBER: US/60/234,509
PRIOR FILING DATE: 2000-09-22
PRIOR APPLICATION NUMBER: US/60/234,567
PRIOR FILING DATE: 2000-09-22
NUMBER OF SEQ ID NOS: 1392
SOFTWARE: Patent version 3.0
SEQ ID NO 348
LENGTH: 18596
TYPE: DNA
ORGANISM: Homo sapiens
US-09-954-531-348

Query Match 31.7%; Score 487; DB 9; Length 18596;

Best Local Similarity 99.7%; Pred. No. 4,5e-249; Indels 0; Gaps 0;

Matches 587; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 908 AGCTTCAGGAGAACCCAGACCTTTCCAAAGCTCAGAGATCTTGGAAATTTGAGAAA 967
DB 15611 AGCTTCAGGAGAACCCAGACCTTTCCAAAGCTCAGAGATCTTGGAAATTTGAGAAA 15670
QY 968 TTGATGACTTCAAGCTGAAGCTTTGAGATGAAGGTACATCCGATCAACTTTA 1027
DB 15671 TTGATGACTTCAAGCTGAAGCTTTGAGATGAAGGTACATCCGATCAACTTTA 15730
QY 1028 AAATGAAATGAGCTGTTAGGAGCTTTCAAGAGCTGAAAGATATTTGCACTTTA 1087
DB 15731 AAATGAAATGAGCTGTTAGGAGCTTTCAAGAGCTGAAAGATATTTGCACTTTA 15790
QY 1088 GGGGTTGGGCTGAGATCCGAGGTAAAGTTCTTTTCTTAAAGAAAGAACTAGG 1147
DB 15791 GGGGTTGGGCTGAGATCCGAGGTAAAGTTCTTTTCTTAAAGAAAGAACTAGG 15850
QY 1148 TCAAAAATCTGTCGATGACATGATTAATTTTAAAGATGTTGCACTGGCAAT 1207

DB 15851 TCAAAATCTGTCCGTGACCTTATGATTAATTTTAAAGATGTGCCACTGCAAAAT 15910
QY 1208 GAACTGTGCGAGTCTTCCATAATTAAGAGCTTTGAGTACTCACTGAGGATATCTG 1267
DB 15911 GTAACTGTGCGAGTCTTCCATAATTAAGAGCTTTGAGTACTCACTGAGGATATCTG 15970
QY 1268 ACAATGCTGAGGTTATGAACAAAGTGAAGAGATGAATGTATGTGCTTTAGCAAAAC 1327
DB 15971 ACAATGCTGAGGTTATGAACAAAGTGAAGAGATGAATGTATGTGCTTTAGCAAAAC 16030
QY 1328 ATGTATGTGCAATTTCAATCCACGACTTTAAAGAGTGTGTAATTTCAACACTAT 1387
DB 16031 ATGTATGTGCAATTTCAATCCACGACTTTAAAGAGTGTGTAATTTCAACACTAT 16090
QY 1388 TTTTGGAAATTTTAAATTTTAAAGATTTCAACACTATCCCTCAATCTGAGG 1447
DB 16091 TTTTGGAAATTTTAAATTTTAAAGATTTCAACACTATCCCTCAATCTGAGG 16150
QY 1448 AGCTGATGAACCAATCATCATCATCATCATCATCATCATCATCATCATCATCAT 1496
DB 16151 AGCTGATGAACCAATCATCATCATCATCATCATCATCATCATCATCATCATCAT 16199

RESULT 4
US-09-880-107-1590
; Sequence 1590, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-MO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1590
; LENGTH: 18596
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 D00596
US-09-880-107-1590

Query Match 31.7%; Score 487; DB 10; Length 18596;
Best Local Similarity 99.7%; Pred. No. 4.5e-249;
Matches 587; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 908 AGCTTCAAGGAGAACCCAGACCTTTCCAAAGCTCAGGATTTCTGAAAAGTTGAGAAA 967
DB 15611 AGCTTCAAGGAGAACCCAGACCTTTCCAAAGCTCAGGATTTCTGAAAAGTTGAGAAA 15670
QY 968 TTGATGACTTCAAGCTGAAGACTTTCAGATTTGAAGGGTCAATCCGATCAACTATTA 1027
DB 15671 TTGATGACTTCAAGCTGAAGACTTTCAGATTTGAAGGGTCAATCCGATCAACTATTA 15730
QY 1028 AAAAGGAAAGGCTGTTAGGAGTCTTCAAGAGCTGAGAGATTTGTCAGTCTTA 1087
DB 15731 AAAAGGAAAGGCTGTTAGGAGTCTTCAAGAGCTGAGAGATTTGTCAGTCTTA 15790
QY 1088 GGGGTGGGCTGATGCGAGGTAAGATTTCTTTTGTCTTAAGAAAGAAAGAACTAGG 1147
DB 15791 GGGGTGGGCTGATGCGAGGTAAGATTTCTTTTGTCTTAAGAAAGAAAGAACTAGG 15850
QY 1148 TCAAAATCTGTCCGTGACCTTATGATTAATTTTAAAGATGTGCCACTGCAAAAT 1207

DB 15851 TCAAAATCTGTCCGTGACCTTATGATTAATTTTAAAGATGTGCCACTGCAAAAT 15910
QY 1208 GAACTGTGCGAGTCTTCCATAATTAAGAGCTTTGAGTACTCACTGAGGATATCTG 1267
DB 15911 GTAACTGTGCGAGTCTTCCATAATTAAGAGCTTTGAGTACTCACTGAGGATATCTG 15970
QY 1268 ACAATGCTGAGGTTATGAACAAAGTGAAGAGATGAATGTATGTGCTTTGCAAAAC 1327
DB 15971 ACAATGCTGAGGTTATGAACAAAGTGAAGAGATGAATGTATGTGCTTTGCAAAAC 16030
QY 1328 ATGTATGTGCAATTTCAATCCACGACTTTAAAGAGTGTGTAATTTCAACACTAT 1387
DB 16031 ATGTATGTGCAATTTCAATCCACGACTTTAAAGAGTGTGTAATTTCAACACTAT 16090
QY 1388 TTTTGGAAATTTTAAATTTTAAAGATTTCAACACTATCCCTCAATCTGAGG 1447
DB 16091 TTTTGGAAATTTTAAATTTTAAAGATTTCAACACTATCCCTCAATCTGAGG 16150
QY 1448 AGCTGATGAACCAATCATCATCATCATCATCATCATCATCATCATCATCATCAT 1496
DB 16151 AGCTGATGAACCAATCATCATCATCATCATCATCATCATCATCATCATCATCAT 16199

RESULT 5
US-09-967-768A-119
; Sequence 119, Application US/09967768A
; Patent No. US20020150877A1
; GENERAL INFORMATION:
; APPLICANT: Augustus, Meena
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-72
; CURRENT APPLICATION NUMBER: US/09/967,768A
; CURRENT FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US/60/236,109
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,034
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,111
; PRIOR FILING DATE: 2000-09-28
; NUMBER OF SEQ ID NOS: 325
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 119
; LENGTH: 18596
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-967-768A-119

Query Match 31.7%; Score 487; DB 10; Length 18596;
Best Local Similarity 99.7%; Pred. No. 4.5e-249;
Matches 587; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 908 AGCTTCAAGGAGAACCCAGACCTTTCCAAAGCTCAGGATTTCTGAAAAGTTGAGAAA 967
DB 15611 AGCTTCAAGGAGAACCCAGACCTTTCCAAAGCTCAGGATTTCTGAAAAGTTGAGAAA 15670
QY 968 TTGATGACTTCAAGCTGAAGACTTTCAGATTTGAAGGGTCAATCCGATCAACTATTA 1027
DB 15671 TTGATGACTTCAAGCTGAAGACTTTCAGATTTGAAGGGTCAATCCGATCAACTATTA 15730
QY 1028 AAAAGGAAAGGCTGTTAGGAGTCTTCAAGAGCTGAGAGATTTGTCAGTCTTA 1087
DB 15731 AAAAGGAAAGGCTGTTAGGAGTCTTCAAGAGCTGAGAGATTTGTCAGTCTTA 15790
QY 1088 GGGGTGGGCTGATGCGAGGTAAGATTTCTTTTGTCTTAAGAAAGAAAGAACTAGG 1147
DB 15791 GGGGTGGGCTGATGCGAGGTAAGATTTCTTTTGTCTTAAGAAAGAAAGAACTAGG 15850
QY 1148 TCAAAATCTGTCCGTGACCTTATGATTAATTTTAAAGATGTGCCACTGCAAAAT 1207
DB 15851 TCAAAATCTGTCCGTGACCTTATGATTAATTTTAAAGATGTGCCACTGCAAAAT 15910
QY 1208 GTAACTGTGCGAGTCTTCCATAATTAAGAGCTTTGAGTACTCACTGAGGATATCTG 1267

Db 15911 GTACCTGTCAGGTTCTTCCATTAATTAAGCTTTGAGTTAACTCACTGAGGATATCTG 15970
QY 1268 ACAATGCTAGGTTATGAACAAAGTAGAGAAATGAATGATATGCTCTTACCAAAAAC 1327
Db 15971 ACAATGCTAGGTTATGAACAAAGTAGAGAAATGAATGATATGCTCTTACCAAAAAC 16030
QY 1328 ATGTATGCTATTCATCCAGCTATCTATTAAGAGTTGGTGAATTTCCCAAGCTAT 1387
Db 16031 ATGTATGCTATTCATCCAGCTATCTATTAAGAGTTGGTGAATTTCCCAAGCTAT 16090
QY 1388 TTTTGAATATTTTATGAATTTTATGAATTTTCAAGCTATTTCCCAATCTGAGG 1447
Db 16091 TTTTGAATATTTTATGAATTTTATGAATTTTCAAGCTATTTCCCAATCTGAGG 16150
QY 1448 AGCTGATACACATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 1496
Db 16151 AGCTGATACACATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 16199

RESULT 6

US-09-922-217-443/c
; Sequence 443, Application US/09922217
; Patent No. US20020076414A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tonglong
; APPLICANT: Jiang, Yugu
; APPLICANT: Smith, Carole Lynn
; APPLICANT: King, Gordon E.
; APPLICANT: Wang, Aljun
; APPLICANT: Clapper, Jonathan D.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; TITLE OF INVENTION: OF COLON CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.471C13
; CURRENT APPLICATION NUMBER: US/09/922,217
; CURRENT FILING DATE: 2001-08-03
; NUMBER OF SEQ ID NOS: 1124
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 443
; LENGTH: 346
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-922-217-443

Query Match

Best Local Similarity 98.8%; Score 142; DB 10; Length 346;
Matches 342; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1007 ACAATCCGATCCATCTATTAATGAATGCTCTTTAGGCTTTTCAAGAGCTN 1066
Db 346 ACAATCCGATCCATCTATTAATGAATGCTCTTTAGGCTTTTCAAGAGCTT 287
QY 1067 GAAGGATATTCAGTCTTTAGGGCTTTGGGCTGGATCCGAGTAAAGTTCTTTTGT 1126
Db 286 GAAGGATATTCAGTCTTTAGGGCTTTGGGCTGGATCCGAGTAAAGTTCTTTTGT 227
QY 1127 CTAAGAANAAGAACTAGTCAAAATCTGCTGACATCATCATCTATTAATTTTAA 1186
Db 226 CTAAGAANAAGAACTAGTCAAAATCTGCTGACATCATCATCTATTAATTTTAA 167
QY 1187 AGGATGTTCCCATCTGGCAATGTAATCTGCTGACATCATCTATTAATTTTAA 1246
Db 166 AGGATGTTCCCATCTGGCAATGTAATCTGCTGACATCATCTATTAATTTTAA 107
QY 1247 TTAATCACTGAGGTTATCTGCAATGCTGATGTAATGTAATGTAATGTAAT 1306
Db 106 TTAATCACTGAGGTTATCTGCAATGCTGATGTAATGTAATGTAATGTAAT 47

QY 1307 GTATGCTCTTAGCAAAAACATGATGCTATTTGATCCACGT 1352
Db 46 GTATGCTCTTAGCAAAAACATGATGCTATTTGATCCACGT 1

RESULT 7

US-09-833-263-443/c
; Sequence 443, Application US/09833263
; Patent No. US20020110547A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Aljun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Stolk, John A.
; APPLICANT: Meagher, Madeleine J.
; TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.471C12
; CURRENT APPLICATION NUMBER: US/09/833,263
; CURRENT FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 1093
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 443
; LENGTH: 346
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-833-263-443

Query Match

Best Local Similarity 98.8%; Score 142; DB 10; Length 346;
Matches 342; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1007 ACAATCCGATCCATCTATTAATGAATGCTCTTTAGGCTTTTCAAGAGCTN 1066
Db 346 ACAATCCGATCCATCTATTAATGAATGCTCTTTAGGCTTTTCAAGAGCTT 287
QY 1067 GAAGGATATTCAGTCTTTAGGGCTTTGGGCTGGATCCGAGTAAAGTTCTTTTGT 1126
Db 286 GAAGGATATTCAGTCTTTAGGGCTTTGGGCTGGATCCGAGTAAAGTTCTTTTGT 227
QY 1127 CTAAGAANAAGAACTAGTCAAAATCTGCTGACATCATCATCTATTAATTTTAA 1186
Db 226 CTAAGAANAAGAACTAGTCAAAATCTGCTGACATCATCATCTATTAATTTTAA 167
QY 1187 AGGATGTTCCCATCTGGCAATGTAATCTGCTGACATCATCTATTAATTTTAA 1246
Db 166 AGGATGTTCCCATCTGGCAATGTAATCTGCTGACATCATCTATTAATTTTAA 107
QY 1247 TTAATCACTGAGGTTATCTGCAATGCTGATGTAATGTAATGTAATGTAAT 1306
Db 106 TTAATCACTGAGGTTATCTGCAATGCTGATGTAATGTAATGTAATGTAAT 47
QY 1307 GTATGCTCTTAGCAAAAACATGATGCTATTTGATCCACGT 1352
Db 46 GTATGCTCTTAGCAAAAACATGATGCTATTTGATCCACGT 1

RESULT 8

US-09-797-163-2
; Sequence 2, Application US/09797163
; Patent No. US2002009795A1
; GENERAL INFORMATION:
; APPLICANT: University of Southern California
; APPLICANT: Danenberg, Kathleen D.
; APPLICANT: Danenberg, Peter V.
; APPLICANT: Swenson, Steven
; TITLE OF INVENTION: ISOLATION OF RNA, DNA AND PROTEINS FROM
; TITLE OF INVENTION: FORMALIN-FIXED PARAFFIN-EMBEDDED TISSUE SPECIMENS
; FILE REFERENCE: 13761-770
; CURRENT APPLICATION NUMBER: US/09/797,163
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: US 09/469,338
; PRIOR FILING DATE: 1999-12-20

```
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 22
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: PCR primer
US-09-797-163-2

Query Match          1.4%; Score 22; DB 10; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.093;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 781 AACATGCGAGCTAGCCCTGC 802
DB 1 AACATGCGAGCTAGCCCTGC 22

RESULT 9
US-09-998-333A-9
; Sequence 9, Application US/09998333A
; Patent No. US20020142328A1
; GENERAL INFORMATION:
; APPLICANT: Kathleen D. Danenberg
; TITLE OF INVENTION: METHOD OF DETERMINING A CHEMOTHERAPEUTIC
; TITLE OF INVENTION: REGIMEN BY ASSAYING GENE EXPRESSION IN PRIMARY TUMORS
; FILE REFERENCE: 11220/146
; CURRENT APPLICATION NUMBER: US/09/998,333A
; PRIOR FILING DATE: 2001-12-03
; PRIOR APPLICATION NUMBER: To Be Assigned
; PRIOR FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 09/879,217
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: 09/877,178
; PRIOR FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: 09/877,177
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: 60/250,472
; PRIOR FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: 60/250,120
; PRIOR FILING DATE: 2000-12-01
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 22
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Oligonucleotide Primer
US-09-998-333A-9

Query Match          1.4%; Score 22; DB 10; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.093;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 781 AACATGCGAGCTAGCCCTGC 802
DB 1 AACATGCGAGCTAGCCCTGC 22

RESULT 10
US-09-797-163-3/c
; Sequence 3, Application US/09797163
; Patent No. US2002009795A1
; GENERAL INFORMATION:
; APPLICANT: University of Southern California
; APPLICANT: Danenberg, Kathleen D.
; APPLICANT: Danenberg, Peter V.
; APPLICANT: Swenson, Steven
; TITLE OF INVENTION: ISOLATION OF RNA, DNA AND PROTEINS FROM
; TITLE OF INVENTION: FORMALIN-FIXED PARAFFIN-EMBEDDED TISSUE SPECIMENS
; FILE REFERENCE: 13761-770
```

```
; CURRENT APPLICATION NUMBER: US/09/797,163
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: US 09/469,338
; PRIOR FILING DATE: 1999-12-20
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 21
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: PCR primer
US-09-797-163-3

Query Match          1.4%; Score 21; DB 10; Length 21;
Best Local Similarity 100.0%; Pred. No. 0.31;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 805 ACGTACATGATTGGCGCATC 825
DB 21 ACGTACATGATTGGCGCATC 1

RESULT 11
US-09-998-333A-8/c
; Sequence 8, Application US/09998333A
; Patent No. US20020142328A1
; GENERAL INFORMATION:
; APPLICANT: Kathleen D. Danenberg
; TITLE OF INVENTION: METHOD OF DETERMINING A CHEMOTHERAPEUTIC
; TITLE OF INVENTION: REGIMEN BY ASSAYING GENE EXPRESSION IN PRIMARY TUMORS
; FILE REFERENCE: 11220/146
; CURRENT APPLICATION NUMBER: US/09/998,333A
; PRIOR FILING DATE: 2001-12-03
; PRIOR APPLICATION NUMBER: To Be Assigned
; PRIOR FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 09/879,217
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: 09/877,178
; PRIOR FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: 09/877,177
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: 60/250,472
; PRIOR FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: 60/250,120
; PRIOR FILING DATE: 2000-12-01
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 21
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Oligonucleotide Primer
US-09-998-333A-8

Query Match          1.4%; Score 21; DB 10; Length 21;
Best Local Similarity 100.0%; Pred. No. 0.31;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 805 ACGTACATGATTGGCGCATC 825
DB 21 ACGTACATGATTGGCGCATC 1

RESULT 12
US-09-822-485-34
; Sequence 34, Application US/09822485
; Patent No. US20020001825A1
; GENERAL INFORMATION:
; APPLICANT: Itoh, No. US20020001825A1uyuki
; TITLE OF INVENTION: No. US20020001825A1el Fibroblast Growth Factor-Like Polypeptides
; FILE REFERENCE: 08035.0001-01000
```

;; CURRENT APPLICATION NUMBER: US/09/822,485
;; CURRENT FILING DATE: 2001-04-02
;; NUMBER OF SEQ ID NOS: 35
;; SOFTWARE: Patentin Ver. 2.0
;; SEQ ID NO 34
;; LENGTH: 995
;; TYPE: DNA
;; ORGANISM: Mus sp.
US-09-822-485-34

Query Match 1.3%; Score 20; DB 10; Length 995;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGGGGGGGGGACCACTTG 20
DB 39 GGGGGGGGGGGGACCACTTG 58

RESULT 13
US-09-954-197-1
; Sequence 1, Application US/09954197
; Patent No. US20020107379A1
; GENERAL INFORMATION:
; APPLICANT: MAX, Achim
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES CODING FOR THE CHYA GENE
; FILE REFERENCE: 032301 MD 201
; CURRENT APPLICATION NUMBER: US/09/954,197
; CURRENT FILING DATE: 2001-12-18
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 1
; LENGTH: 1200
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (231)..(1028)
; OTHER INFORMATION:
US-09-954-197-1

Query Match 1.3%; Score 20; DB 10; Length 1200;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 769 GGTGTGCTTTCACATGCG 788
DB 756 GGTGTGCTTTCACATGCG 775

RESULT 14
US-09-987-446-1
; Sequence 1, Application US/09987446
; Patent No. US20020115162A1
; GENERAL INFORMATION:
; APPLICANT: FARWICK, Mike, et al.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES CODING FOR THE CYSQ GENE
; FILE REFERENCE: 032301 MD 245
; CURRENT APPLICATION NUMBER: US/09/987,446
; CURRENT FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 1
; LENGTH: 2730
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1014)..(1769)
; OTHER INFORMATION:
US-09-987-446-1

Query Match 1.3%; Score 20; DB 10; Length 2730;

Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 769 GGTGTGCTTTCACATGCG 788
DB 2316 GGTGTGCTTTCACATGCG 2335

RESULT 15
US-09-815-343-185/C
; Sequence 185, Application US/09815343
; Patent No. US20010055596A1
; GENERAL INFORMATION:
; APPLICANT: Meagher, Madeleine
; APPLICANT: Xu, Jiangchun
; APPLICANT: King, Gordon B.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.504
; CURRENT APPLICATION NUMBER: US/09/815,343
; CURRENT FILING DATE: 2001-03-22
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 185
; LENGTH: 358
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(358)
; OTHER INFORMATION: n = A,T,C or G
US-09-815-343-185

Query Match 1.2%; Score 19; DB 10; Length 358;
Best Local Similarity 100.0%; Pred. No. 4.5;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1430 TTCCCTGAATCTGAGGGA 1448
DB 297 TTCCCTGAATCTGAGGGA 279

Search completed: November 29, 2002, 05:48:25
Job time : 631.315 secs

PAGES: 2035-2043
DATE: 25-MAR-1985
US-09-089-195-1

Query Match	90.8%;	Score 1394;	DB 3;	Length 1536;
Best Local Similarity	99.9%;	Pred. No. 0;		
Matches 1494;	Conservative	0;	Mismatches 2;	Indels 0;
				Gaps 0.

[illegible]

Db	961	GAGAAATTCGACCTTCAAGCTGAGACTTTGACATTGAAAGGTACATCCGATCCA	1020
Qy	1021	ACTATTAATGGAATGCGCTGTTTAGGTCTCTCAAGAGAGCTNAGATATGTCA	1080
Db	1021	ACTATTAATGGAATGCGCTGTTTAGGTCTCTCAAGAGAGCTNAGATATGTCA	1080
Qy	1081	GTCTTAGGGGTGGGGCTGGAATCCGAGGTAAAGTCTTTTGCTTAAAGANAAAG	1140
Db	1081	GTCTTAGGGGTGGGGCTGGAATCCGAGGTAAAGTCTTTTGCTTAAAGANAAAG	1140
Qy	1141	AACTAGGTCAAAAATCTGTCCGTGACCTATGATTAATTTTAAAGATGTTCACCT	1200
Db	1141	AACTAGGTCAAAAATCTGTCCGTGACCTATGATTAATTTTAAAGATGTTCACCT	1200
Qy	1201	GGCAATATGATCTGTGCAGTCTTTCCTCAATATAAAGGCTTTAGATTACTCAGAG	1260
Db	1201	GGCAATATGATCTGTGCAGTCTTTCCTCAATATAAAGGCTTTAGATTACTCAGAG	1260
Qy	1261	GTATCTGACATCTGAGGTTATGACAAGAGGAGGAGATGAATGTATGCTCTAG	1320
Db	1261	GTATCTGACATCTGAGGTTATGACAAGAGGAGGAGATGAATGTATGCTCTAG	1320
Qy	1321	CAAAAATGATATGTGCTCAATCCACGATCTAAAGAGGTTGATTTTCAAC	1380
Db	1321	CAAAAATGATATGTGCTCAATCCACGATCTAAAGAGGTTGATTTTCAAC	1380
Qy	1381	AAGCTATTTTGGATATTTTAAATATTTTAAAGATTTCAAGACTATTCCTCAAT	1440
Db	1381	AAGCTATTTTGGATATTTTAAATATTTTAAAGATTTCAAGACTATTCCTCAAT	1440
Qy	1441	CTGAGGAGCTGATGAACCATCGATCATGATGATGAGGTGTTATGAACTTTA	1496
Db	1441	CTGAGGAGCTGATGAACCATCGATCATGATGATGAGGTGTTATGAACTTTA	1496

```

US-09-367-007C-38
RESULT 2
US-09-367-007C-38
Sequence 38, Application US/09367007C
Patent No. 6416987
GENERAL INFORMATION:
APPLICANT: Bertino, Joseph R.
APPLICANT: Banerjee, Debabrata
APPLICANT: Tong, Youzhi
APPLICANT: Liu-Chen, Xinyue
TITLE OF INVENTION: Mutants of Thymidylate Synthase and Uses Thereof
FILE REFERENCE: D5978
CURRENT APPLICATION NUMBER: US/09367, 007C
CURRENT FILING DATE: 1999-10-15
PRIOR APPLICATION NUMBER: PCT/US98/02145
PRIOR FILING DATE: 1998-01-03
NUMBER OF SEQ ID NOS: 39
SEQ ID NO 38
LENGTH: 1536
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: wild type human thymidylate synthase cDNA
OTHER INFORMATION: (Genbank Accession number IM 001071)
US-09-367-007C-38

```

Query Match	90.8%;	Score 1394;	DB 4;	Length 1536;
Best Local Similarity	99.9%;	Pred. No. 0;	Mismatches 2;	Indels 0;
Matches 1494;	Conservative	0;	0;	0;
QY	1	GGGGGGGGGGGACACTTGGCGTCCGTCGGCGCGGACACTTGGCGTCCGTCGGT	60	
Db	1	GGGGGGGGGGGACACTTGGCGTCCGTCGGTCGGCGCGGACACTTGGCGTCCGTCGGT	60	
QY	61	CCGGCGCGCGCACTTGGCGTCCGTCGGTCGGCGCGGCGCGGCACTGCTGTGGCGCGG	120	
Db	61	CCGGCGCGCGCACTTGGCGTCCGTCGGTCGGTCGGCGCGGCGCGGCACTGCTGTGGCGCGG	120	

QY 526 CATTTCGGGCGAATACAGAGATATGATATTCATTCAGAGACGAGATTGACCA 585
|||
DB 421 CATTTCGGGCGAATACAGAGATATGATATTCATTCAGAGACGAGATTGACCA 480
QY 586 CTGCAAAAGATGATTGACACCATCAAAACCAACCTGACGACAGAGAAATCATCATGTC 645
|||
DB 481 CTGCAAAAGATGATTGACACCATCAAAACCAACCTGACGACAGAGAAATCATCATGTC 540
QY 646 GCTTGGAATCCAGAGATCTTCTCTGATGAGCGCTGCTCCATGCAATGCTCTGCGAG 705
|||
DB 541 GCTTGGAATCCAGAGATCTTCTCTGATGAGCGCTGCTCCATGCAATGCTCTGCGAG 600
QY 706 TTCTATGATGAAACAGTGAAGCTGCTGTCGACGTTTACAGAGATCGGAGACATGAGG 765
|||
DB 601 TTCTATGATGAAACAGTGAAGCTGCTGTCGACGTTTACAGAGATCGGAGACATGAGG 660
QY 766 CTGAGTGTCTTCAACATGCGACGCTGACGCTGCTCAACATGATGATGAGGACATC 825
|||
DB 661 CTGAGTGTCTTCAACATGCGACGCTGACGCTGCTCAACATGATGATGAGGACATC 720
QY 826 AGGAGCTGAGACGAGGATCTTATACACTTTGGAGATGACATTTTACCTGAT 885
|||
DB 721 AGGAGCTGAGACGAGGATCTTATACACTTTGGAGATGACATTTTACCTGAT 780
QY 886 CACATGAGCCATGAAATTCAGCTTCAGCAGAACCCAGACCTTCCAAAGCTCAG 945
|||
DB 781 CACATGAGCCATGAAATTCAGCTTCAGCAGAACCCAGACCTTCCAAAGCTCAG 840
QY 946 ATCTTCGAAAAGTTGAGAAATTTGATGACTTCAAGCTGAAGCTTCAGATTGAAGG 1005
|||
DB 841 ATCTTCGAAAAGTTGAGAAATTTGATGACTTCAAGCTGAAGCTTCAGATTGAAGG 900
QY 1006 TACATTCGATCCATCTATTTAAATGGAATGCTGTTTG 1047
|||
DB 901 TACATTCGATCCATCTATTTAAATGGAATGCTGTTTG 942

RESULT 4
US-09-318-448-11
; Sequence 11, Application US/09318448
; Patent No. 6210950
; GENERAL INFORMATION:
; APPLICANT: Johnson, William G.
; APPLICANT: Stencos, Edward S.
; TITLE OF INVENTION: METHODS FOR DIAGNOSING, PREVENTING, AND TREATING
; FILE OF INVENTION: DEVELOPMENTAL DISORDERS
; FILE REFERENCE: 601-1-057
; CURRENT APPLICATION NUMBER: US/09/318,448
; CURRENT FILING DATE: 1999-05-25
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 18596
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-318-448-11

Query Match 31.7%; Score 487; DB 4; Length 18596;
Best Local Similarity 99.7%; Pred. No. 4e-241;
Matches 587; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 908 AGCTTCAGGAGAACCCAGACCTTCCCAAGCTCAGGATCTCTGAAAAGTTGAAAA 967
|||
DB 15611 AGCTTCAGGAGAACCCAGACCTTCCCAAGCTCAGGATCTCTGAAAAGTTGAAAA 15670
QY 968 TTGATGACTTCAAGCTGAGACTTCAGATTGAAGGATCAATCCGATCCAACTATTA 1027
|||
DB 15671 TTGATGACTTCAAGCTGAGACTTCAGATTGAAGGATCAATCCGATCCAACTATTA 15730
QY 1028 AAATGGAATGCTGTTTAAAGGCTTCAAAAGGAGCTGAAAGATATTTGAGCTTTTA 1087
|||
DB 15731 AAATGGAATGCTGTTTAAAGGCTTCAAAAGGAGCTGAAAGATATTTGAGCTTTTA 15790

QY 1088 GGGGTGGGCTGGATGCGAGGTAATAAGTTCTTTTGTCTTAAAGAAAGAACTAGG 1147
|||
DB 15791 GGGGTGGGCTGGATGCGAGGTAATAAGTTCTTTTGTCTTAAAGAAAGAACTAGG 15850
QY 1148 TCAAAATCTGTCCGATGACCTATCAGTTATTAATTTTAAAGATGTTGCTAGGCAAT 1207
|||
DB 15851 TCAAAATCTGTCCGATGACCTATCAGTTATTAATTTTAAAGATGTTGCTAGGCAAT 15910
QY 1208 GTAACTGTGCAAGTTCTTTCATTAATAAAGCTTTGATTAATCTAGGAGGATGTCG 1267
|||
DB 15911 GTAACTGTGCAAGTTCTTTCATTAATAAAGCTTTGATTAATCTAGGAGGATGTCG 15970
QY 1268 ACATATGCTGAGTTATGAACAAAGTGAAGAAATGAATGATGCTCTTTCGAAAAC 1327
|||
DB 15971 ACATATGCTGAGTTATGAACAAAGTGAAGAAATGAATGATGCTCTTTCGAAAAC 16030
QY 1328 ATGATATGCAATTTCAATCCAGTATCTTATTAAGAGGTTGGATTTCAAGCTAT 1387
|||
DB 16031 ATGATATGCAATTTCAATCCAGTATCTTATTAAGAGGTTGGATTTCAAGCTAT 16090
QY 1388 TTTTGAATATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTT 1447
|||
DB 16091 TTTTGAATATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTTT 16150
QY 1448 AGCTGATGACCATGATCATGATGATGATGATGATGATGATGATGATGATGATGAT 1496
|||
DB 16151 AGCTGATGACCATGATCATGATGATGATGATGATGATGATGATGATGATGATGAT 16199

RESULT 5
US-08-965-048-5/C
; Sequence 5, Application US/08965048
; Patent No. 6323244
; GENERAL INFORMATION:
; APPLICANT: Chen, Hong
; APPLICANT: Freiler, Nelson
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND
; FILE OF INVENTION: TREATMENT OF NEUROPSYCHIATRIC DISORDERS
; FILE REFERENCE: 7853-093
; CURRENT APPLICATION NUMBER: US/08/965,048
; CURRENT FILING DATE: 1997-11-05
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 45716
; TYPE: DNA
; ORGANISM: Homo sapiens
US-08-965-048-5

Query Match 31.7%; Score 487; DB 4; Length 45716;
Best Local Similarity 99.7%; Pred. No. 4.1e-241;
Matches 587; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 908 AGCTTCAGGAGAACCCAGACCTTCCCAAGCTCAGGATCTCTGAAAAGTTGAAAA 967
|||
DB 44582 AGCTTCAGGAGAACCCAGACCTTCCCAAGCTCAGGATCTCTGAAAAGTTGAAAA 44523
QY 968 TTGATGACTTCAAGCTGAGACTTCAGATTGAAGGATCAATCCGATCCAACTATTA 1027
|||
DB 44522 TTGATGACTTCAAGCTGAGACTTCAGATTGAAGGATCAATCCGATCCAACTATTA 44463
QY 1028 AAATGGAATGCTGTTTAAAGGCTTCAAAAGGAGCTGAAAGATATTTGAGCTTTTA 1087
|||
DB 44462 AAATGGAATGCTGTTTAAAGGCTTCAAAAGGAGCTGAAAGATATTTGAGCTTTTA 44403
QY 1088 GGGGTGGGCTGGATGCGAGGTAATAAGTTCTTTTGTCTTAAAGAAAGAACTAGG 1147
|||
DB 44402 GGGGTGGGCTGGATGCGAGGTAATAAGTTCTTTTGTCTTAAAGAAAGAACTAGG 44343
QY 1148 TCAAAATCTGTCCGATGACCTATCAGTTATTAATTTTAAAGATGTTGCTAGGCAAT 1207
|||
DB 44342 TCAAAATCTGTCCGATGACCTATCAGTTATTAATTTTAAAGATGTTGCTAGGCAAT 44283

QY 1208 GTAACGTGCGAAGTCTTCCATATATAAAGCTTTGAGTACTCACTGAGGGTATCTG 1267
 Db 44282 GTAACGTGCGAAGTCTTCCATATATAAAGCTTTGAGTACTCACTGAGGGTATCTG 44223
 QY 1268 ACAATGCTGAGGTTATGAACAAGTGAAGGAAATGAATGATGAGCTTGAAGCAAAAAC 1327
 Db 44222 ACAATGCTGAGGTTATGAACAAGTGAAGGAAATGAATGATGAGCTTGAAGCAAAAAC 44163
 QY 1328 ATGATGCTGATTTCAATCCACGACTTATAAAGAGTGGTAATTTCAACAAGCTAT 1387
 Db 44162 ATGATGCTGATTTCAATCCACGACTTATAAAGAGTGGTAATTTCAACAAGCTAT 44103
 QY 1388 TTTTGGAAATTTTGAATATTTTGAATTTTCAAGACTTATCCCTCAAAATCTGAGG 1447
 Db 44102 TTTTGGAAATTTTGAATATTTTGAATTTTCAAGACTTATCCCTCAAAATCTGAGG 44043
 QY 1448 AGCTGAGTAACCACTGATGCTATGATGAGAGTGTGTATGAATTTA 1436
 Db 44042 AGCTGAGTAACCACTGATGCTATGATGAGAGTGTGTATGAATTTA 43994

RESULT 6

US-08-965-048-6/c
 ; Sequence 6, Application US/08965048
 ; Patent No. 6323244
 ; GENERAL INFORMATION:
 ; APPLICANT: Chen, Hong
 ; APPLICANT: Freeman, Nelson
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND
 ; TREATMENT OF NEUROPSYCHIATRIC DISORDERS
 ; FILE REFERENCE: 7853-093
 ; CURRENT APPLICATION NUMBER: US/08/965,048
 ; CURRENT FILING DATE: 1997-11-05
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 6
 ; LENGTH: 45989
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-08-965-048-6

Query Match 31.7%; Score 487; DB 4; Length 45989;
 Best Local Similarity 99.7%; Pred. No. 4,1e-24;
 Matches 587, Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 908 AGCTTGACGAGAACCCAGACCTTTCCCAAGCTCGAGATTTTGGAAAAGTTGAGAAA 967
 Db 44855 AGCTTGACGAGAACCCAGACCTTTCCCAAGCTCGAGATTTTGGAAAAGTTGAGAAA 44796
 QY 968 TTGATGACTTCAAAAGCTGAAAGCTTTCAGATGAGGTAACATCCGATCCAACTATTA 1027
 Db 44795 TTGATGACTTCAAAAGCTGAAAGCTTTCAGATGAGGTAACATCCGATCCAACTATTA 44736
 QY 1028 AAATGGAATGCTGTTTAAAGGTCCTTCAAGAGCTTGAAGATATGTCACTCTTTA 1087
 Db 44735 AAATGGAATGCTGTTTAAAGGTCCTTCAAGAGCTTGAAGATATGTCACTCTTTA 44676
 QY 1088 GGGGTTGGCTGAGATCCGAGTAAAGTCTTTTGGCTTAAAGATTAAGAAATAGG 1147
 Db 44675 GGGGTTGGCTGAGATCCGAGTAAAGTCTTTTGGCTTAAAGATTAAGAAATAGG 44616
 QY 1148 TCAAAAATCTGTCGAGCTTTCATATAAAGGCTTTGAGTAACTCACTAGGGATCTG 1207
 Db 44615 TCAAAAATCTGTCGAGCTTTCATATAAAGGCTTTGAGTAACTCACTAGGGATCTG 44556
 QY 1208 GTAACGTGCGAAGTCTTCCATATATAAAGCTTTGAGTACTCACTGAGGGTATCTG 1267
 Db 44555 GTAACGTGCGAAGTCTTCCATATATAAAGCTTTGAGTACTCACTGAGGGTATCTG 44496
 QY 1268 ACAATGCTGAGGTTATGAACAAGTGAAGGAAATGAATGATGAGCTTGAAGCAAAAAC 1327
 Db 44495 ACAATGCTGAGGTTATGAACAAGTGAAGGAAATGAATGATGAGCTTGAAGCAAAAAC 44436

QY 1328 ATGATGCTGATTTCAATCCACGACTTATAAAGAGTGGTAATTTCAACAAGCTAT 1387
 Db 44435 ATGATGCTGATTTCAATCCACGACTTATAAAGAGTGGTAATTTCAACAAGCTAT 44376
 QY 1388 TTTTGGAAATTTTGAATATTTTGAATTTTCAAGACTTATCCCTCAAAATCTGAGG 1447
 Db 44375 TTTTGGAAATTTTGAATATTTTGAATTTTCAAGACTTATCCCTCAAAATCTGAGG 44316
 QY 1448 AGCTGAGTAACCACTGATGCTATGATGAGAGTGTGTATGAATTTA 1436
 Db 44315 AGCTGAGTAACCACTGATGCTATGATGAGAGTGTGTATGAATTTA 44267

RESULT 7

US-08-965-048-1/c
 ; Sequence 1, Application US/08965048
 ; Patent No. 6323244
 ; GENERAL INFORMATION:
 ; APPLICANT: Chen, Hong
 ; APPLICANT: Freeman, Nelson
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND
 ; TREATMENT OF NEUROPSYCHIATRIC DISORDERS
 ; FILE REFERENCE: 7853-093
 ; CURRENT APPLICATION NUMBER: US/08/965,048
 ; CURRENT FILING DATE: 1997-11-05
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 1
 ; LENGTH: 1817
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (67)..(1149)
 ; US-08-965-048-1

Query Match 10.9%; Score 168; DB 4; Length 1817;
 Best Local Similarity 99.5%; Pred. No. 8,6e-77;
 Matches 218, Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 908 AGCTTGACGAGAACCCAGACCTTTCCCAAGCTCGAGATTTTGGAAAAGTTGAGAAA 967
 Db 1515 AGCTTGACGAGAACCCAGACCTTTCCCAAGCTCGAGATTTTGGAAAAGTTGAGAAA 1456
 QY 968 TTGATGACTTCAAAAGCTGAAAGCTTTCAGATGAGGTAACATCCGATCCAACTATTA 1027
 Db 1455 TTGATGACTTCAAAAGCTGAAAGCTTTCAGATGAGGTAACATCCGATCCAACTATTA 1396
 QY 1028 AAATGGAATGCTGTTTAAAGGTCCTTCAAGAGCTTGAAGATATGTCACTCTTTA 1087
 Db 1395 AAATGGAATGCTGTTTAAAGGTCCTTCAAGAGCTTGAAGATATGTCACTCTTTA 1336
 QY 1088 GGGGTTGGCTGAGATCCGAGTAAAGTCTTTTGGCT 1126
 Db 1335 GGGGTTGGCTGAGATCCGAGTAAAGTCTTTTGGCT 1297

RESULT 8

US-09-367-007C-7
 ; Sequence 7, Application US/09367007C
 ; Patent No. 6416987
 ; GENERAL INFORMATION:
 ; APPLICANT: Bertino, Joseph R.
 ; APPLICANT: Benzer, Deborah
 ; APPLICANT: Tong, Youzhi
 ; TITLE OF INVENTION: Mutants of Thymidylate Synthase and Uses Thereof
 ; FILE REFERENCE: D5978
 ; CURRENT APPLICATION NUMBER: US/09/367,007C
 ; CURRENT FILING DATE: 1999-10-15
 ; PRIOR APPLICATION NUMBER: PCT/US98/02145
 ; PRIOR FILING DATE: 1998-01-03

NUMBER OF SEQ ID NOS: 39
SEQ ID NO 7
LENGTH: 26
TYPE: DNA
ORGANISM: artificial sequence
FEATURE:
NAME/KEY: primer bind
LOCATION: 594..616
OTHER INFORMATION: Sense primer hts-3a for PCR amplification of part
OTHER INFORMATION: of 890 base pair fragment of human TS gene
OTHER INFORMATION: from nucleotide 50 to the C-terminus
US-09-367-007C-7

Query Match 1.6%; Score 26; DB 4; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 699 CTGCCAGTCTATGTGTGAACAGTG 724
Db 1 CTGCCAGTCTATGTGTGAACAGTG 26

RESULT 9
US-09-089-195-26/C
Sequence 26, Application US/09089195
Patent No. 6087489
GENERAL INFORMATION:
APPLICANT: Dean, Nicholas M.
TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDE MODULATION
TITLE OF INVENTION: OF HUMAN THYMIDYLATE SYNTHASE EXPRESSION
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESSES:
ADDRESSES: Law Offices of Jane Massey Licata
STREET: 66 East Main Street
CITY: Marlton
STATE: NJ
COUNTRY: USA
ZIP: 08053
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE:
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/089,195
FILING DATE: herewith
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Jane Massey Licata
REGISTRATION NUMBER: 32,257
REFERENCE/DOCKET NUMBER: ISPH-0309
TELECOMMUNICATION INFORMATION:
TELEPHONE: (609) 810-1515
TELEFAX: (609) 810-1454
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 25 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Linear
US-09-089-195-26

Query Match 1.6%; Score 25; DB 3; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.0036;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1036 ATGGCTGTTAGGCTGCTTCAAG 1060
Db 25 ATGGCTGTTAGGCTGCTTCAAG 1

RESULT 10
US-09-367-007C-11

Sequence 11, Application US/09367007C
Patent No. 6416987
GENERAL INFORMATION:
APPLICANT: Bertino, Joseph R.
APPLICANT: Banerjee, Debabrata
APPLICANT: Tong, Youzhi
APPLICANT: Liu-Chen, Xinyue
TITLE OF INVENTION: Mutants of Thymidylate Synthase and Uses Thereof
FILE REFERENCE: D5978
CURRENT APPLICATION NUMBER: US/09/367,007C
CURRENT FILING DATE: 1999-10-15
PRIOR APPLICATION NUMBER: PCT/US98/02145
PRIOR FILING DATE: 1998-01-03
NUMBER OF SEQ ID NOS: 39
SEQ ID NO 11
LENGTH: 25
TYPE: DNA
ORGANISM: artificial sequence
FEATURE:
NAME/KEY: primer bind
LOCATION: 451..475
OTHER INFORMATION: Sense primer hts-5a for PCR amplification of part
OTHER INFORMATION: of 890 base pair fragment of human TS gene
OTHER INFORMATION: from nucleotide 50 to the C-terminus
US-09-367-007C-11

Query Match 1.6%; Score 25; DB 4; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.0036;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 556 TCAGATTATTCAGACAGGAGGTTG 580
Db 1 TCAGATTATTCAGACAGGAGGTTG 25

RESULT 11
US-09-367-007C-6/C
Sequence 6, Application US/09367007C
Patent No. 6416987
GENERAL INFORMATION:
APPLICANT: Bertino, Joseph R.
APPLICANT: Banerjee, Debabrata
APPLICANT: Tong, Youzhi
APPLICANT: Liu-Chen, Xinyue
TITLE OF INVENTION: Mutants of Thymidylate Synthase and Uses Thereof
FILE REFERENCE: D5978
CURRENT APPLICATION NUMBER: US/09/367,007C
CURRENT FILING DATE: 1999-10-15
PRIOR APPLICATION NUMBER: PCT/US98/02145
PRIOR FILING DATE: 1998-01-03
NUMBER OF SEQ ID NOS: 39
SEQ ID NO 6
LENGTH: 24
TYPE: DNA
ORGANISM: artificial sequence
FEATURE:
NAME/KEY: primer bind
LOCATION: 645..672
OTHER INFORMATION: Antisense primer hts-2b for PCR amplification of part
OTHER INFORMATION: of 890 base pair fragment of human TS gene
OTHER INFORMATION: from nucleotide 50 to the C-terminus
US-09-367-007C-6

Query Match 1.6%; Score 24; DB 4; Length 24;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 727 CTGTCTGCCAGCTGTACACAGAG 750
Db 24 CTGTCTGCCAGCTGTACACAGAG 1

RESULT 12

US-09-367-007C-8/C
; Sequence 8, Application US/09367007C
; Patent No. 6416987

; GENERAL INFORMATION:
; APPLICANT: Bertino, Joseph R.
; APPLICANT: Banerjee, Debabrata
; APPLICANT: Tong, Youzhi
; APPLICANT: Liu-Chen, Xinyue
; TITLE OF INVENTION: Mutants of Thymidylate Synthase and Uses Thereof
; FILE REFERENCE: D5978
; CURRENT APPLICATION NUMBER: US/09/367,007C
; CURRENT FILING DATE: 1999-10-15
; PRIOR APPLICATION NUMBER: PCT/US98/02145
; PRIOR FILING DATE: 1998-01-03
; NUMBER OF SEQ ID NOS: 39
; SEQ ID NO 8
; LENGTH: 24
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
; NAME/KEY: primer bind
; LOCATION: 939..915
; OTHER INFORMATION: Antisense primer hTS-3B for PCR amplification of part
; OTHER INFORMATION: of 890 base pair fragment of human TS gene
; OTHER INFORMATION: from nucleotide 50 to the C-terminus
US-09-367-007C-8

Query Match 1.6%; Score 24; DB 4; Length 24;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1021 ACTATTAATGGAATGCGCTGTT 1044

DB 24 ACTATTAATGGAATGCGCTGTT 1

RESULT 13
US-09-367-007C-13
; Sequence 13, Application US/09367007C
; Patent No. 6416987

; GENERAL INFORMATION:
; APPLICANT: Bertino, Joseph R.
; APPLICANT: Banerjee, Debabrata
; APPLICANT: Tong, Youzhi
; APPLICANT: Liu-Chen, Xinyue
; TITLE OF INVENTION: Mutants of Thymidylate Synthase and Uses Thereof
; FILE REFERENCE: D5978
; CURRENT APPLICATION NUMBER: US/09/367,007C
; CURRENT FILING DATE: 1999-10-15
; PRIOR APPLICATION NUMBER: PCT/US98/02145
; PRIOR FILING DATE: 1998-01-03
; NUMBER OF SEQ ID NOS: 39
; SEQ ID NO 13
; LENGTH: 24
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
; NAME/KEY: primer bind
; LOCATION: 756..779
; OTHER INFORMATION: Sense primer hTS-6A for PCR amplification of part
; OTHER INFORMATION: of 890 base pair fragment of human TS gene
; OTHER INFORMATION: from nucleotide 50 to the C-terminus
US-09-367-007C-13

Query Match 1.6%; Score 24; DB 4; Length 24;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 861 GGGAGATGACATATTACCTGAA 884

DB 1 GGGAGATGACATATTACCTGAA 24

RESULT 14
US-09-367-007C-10/C
; Sequence 10, Application US/09367007C
; Patent No. 6416987

; GENERAL INFORMATION:
; APPLICANT: Bertino, Joseph R.
; APPLICANT: Banerjee, Debabrata
; APPLICANT: Tong, Youzhi
; APPLICANT: Liu-Chen, Xinyue
; TITLE OF INVENTION: Mutants of Thymidylate Synthase and Uses Thereof
; FILE REFERENCE: D5978
; CURRENT APPLICATION NUMBER: US/09/367,007C
; CURRENT FILING DATE: 1999-10-15
; PRIOR APPLICATION NUMBER: PCT/US98/02145
; PRIOR FILING DATE: 1998-01-03
; NUMBER OF SEQ ID NOS: 39
; SEQ ID NO 10
; LENGTH: 23
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
; NAME/KEY: primer bind
; LOCATION: 210..188
; OTHER INFORMATION: Antisense primer hTS-4B for PCR amplification of part
; OTHER INFORMATION: of 890 base pair fragment of human TS gene
; OTHER INFORMATION: from nucleotide 50 to the C-terminus
US-09-367-007C-10

Query Match 1.5%; Score 23; DB 4; Length 23;
Best Local Similarity 100.0%; Pred. No. 0.039;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 293 CCGGCTACAGCTGAGAGATGAA 315

DB 23 CCGGCTACAGCTGAGAGATGAA 1

RESULT 15
US-09-367-007C-12/C
; Sequence 12, Application US/09367007C
; Patent No. 6416987

; GENERAL INFORMATION:
; APPLICANT: Bertino, Joseph R.
; APPLICANT: Banerjee, Debabrata
; APPLICANT: Tong, Youzhi
; APPLICANT: Liu-Chen, Xinyue
; TITLE OF INVENTION: Mutants of Thymidylate Synthase and Uses Thereof
; FILE REFERENCE: D5978
; CURRENT APPLICATION NUMBER: US/09/367,007C
; CURRENT FILING DATE: 1999-10-15
; PRIOR APPLICATION NUMBER: PCT/US98/02145
; PRIOR FILING DATE: 1998-01-03
; NUMBER OF SEQ ID NOS: 39
; SEQ ID NO 12
; LENGTH: 23
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
; NAME/KEY: primer bind
; LOCATION: 503..481
; OTHER INFORMATION: Antisense primer hTS-5B for PCR amplification of part
; OTHER INFORMATION: of 890 base pair fragment of human TS gene
; OTHER INFORMATION: from nucleotide 50 to the C-terminus
US-09-367-007C-12

Query Match 1.5%; Score 23; DB 4; Length 23;
Best Local Similarity 100.0%; Pred. No. 0.039;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 586 CTGCAAGATGATGACACAT 608

DB 23 CTGCAAGATGATGACACAT 1

Sun Dec 1 18:15:47 2002

us-09-963-333-6.oligo.rni

Page 8

Search completed: November 29, 2002, 03:30:35
Job time : 100.613 secs

SUMMARIES

SUMMARIES

Run on: November 28, 2002, 22:48:55 ; Search time 3909.66 Seconds
(without arguments)

11433.695 Million cell updates/sec

Title: US-09-963-333-6

Sequence: 1 ggaggggggggaccactg.....ataataagagtggtctgc 1536

Scoring table: OLIGO_NUC

Searched: 2054640 seqs, 14551402878 residues

Word size : 15

Total number of hits satisfying chosen parameters: 53954

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Listing first 45 summaries

Database ::

1:	gb_ba.*
2:	gb_hng.*
3:	gb_in.*
4:	gb_om.*
5:	gb_ov.*
6:	gb_pac.*
7:	gb_ph.*
8:	gb_pl.*
9:	gb_pr.*
10:	gb_ro.*
11:	gb_scs.*
12:	gb_sy.*
13:	gb_un.*
14:	gb_vl.*
15:	em_ba.*
16:	em_fun.*
17:	em_hum.*
18:	em_in.*
19:	em_mn.*
20:	em_om.*
21:	em_or.*
22:	em_ov.*
23:	em_pat.*
24:	em_pl.*
25:	em_pl.*
26:	em_ro.*
27:	em_scs.*
28:	em_un.*
29:	em_vl.*
30:	em_hng_hum.*
31:	em_hng_inv.*
32:	em_hng_other.*
33:	em_hng_mus.*
34:	em_hng_pln.*
35:	em_hng_rod.*
36:	em_hng_mam.*
37:	em_hng_vrt.*
38:	em_sy.*
39:	em_hngo_hum.*
40:	em_hngo_mus.*
41:	em_hngo_other.*

Pred. No. is the number of results predicted by chance to have a

Result No.	Score	Query Length	DB	ID	Description
1	1394	50.8	1556	6	ARI03889
2	1394	90.8	1556	6	AX397715
3	1394	90.8	1556	6	AX482020
4	1394	90.8	1556	6	HSTXSYN
5	1358	88.4	1563	9	BC013919
6	1357	88.3	1553	9	BC002567
7	1346	87.6	3288	6	AX281791
8	1189	77.4	1567	9	AB062365
9	942	61.3	942	6	AX069335
10	942	61.3	942	6	AX397717
11	532	34.6	555	6	AX341901
12	490	31.9	653	9	AB077207
13	487	31.7	18556	6	ARI44965
14	487	31.7	18556	6	AX050451
15	487	31.7	18556	6	AX330682
16	487	31.7	18556	6	AX330906
17	487	31.7	18556	6	AX357555
18	487	31.7	18556	6	AX397714
19	487	31.7	18556	6	AX408943
20	487	31.7	18556	6	HMTS1
21	487	31.7	43507	9	AF35057
22	487	31.7	152711	9	AF001178
23	487	31.7	159974	2	AC021474
24	487	31.7	165742	2	AC020697
25	456	29.7	840	9	AB077208
26	379	24.7	489	11	G07105
27	297	19.3	1186	9	HMTS
28	274	17.8	165742	2	AC020697
29	168	10.9	1817	9	HSTXSYN
30	150	9.8	196	6	AX135281
31	142	9.2	346	6	AX192876
32	132	8.6	150700	2	AF001097
33	117	7.6	2388	9	AX093873
34	111	7.2	350	11	G14813
35	111	7.2	350	11	G14563
36	111	7.2	1133	6	AX397748
37	107	7.0	860	9	BC028301
38	107	7.0	1131	6	AX397698
39	107	7.0	1161	6	AX397648
40	106	6.9	168	9	AF127520
41	85	5.5	249	9	AF278905
42	82	5.3	140	9	AF127519
43	62	4.0	165	9	AF279907
44	62	4.0	335	9	AF134215
45	59	3.8	196	6	AX135264

ALIGNMENTS

RESULT 1				
AR103889				
LOCUS	AR103889	1536 bp	DNA	linear PAT 14-FEB-2001
DEFINITION	Sequence 1 from patent US 6087489.			
ACCESSION	AR103889			
VERSION	AR103889.1 GI:12815477			
KEYWORDS	.			
SOURCE	Unknown.			
ORGANISM	Unknown.			
REFERENCE	Unclassified.			
AUTHORS	1 (bases 1 to 1536)			
TITLE	Dean, N.M.			
JOURNAL	Anti-sense oligonucleotide modulation of human thymidylate synthase expression Patent : US 6087489-A 1 11-JUL-2000;			

Qy	361	GAGAGTCTGCTGTGTTTATTCAGAGGATCCAAATGTCTAAAGAGCTGCTTCCAAAGGGA	420
Db	361	GAGGAGCTGTGCTGTGTTATTCAGAGGATCCAAATGTCTAAAGAGCTGCTTCCAAAGGA	420
Qy	421	GTCAAAATCTGGGAGGCCAATGAGTCCGAGACTTTTGTGACAGCTCGGATTTCCACC	480
Db	421	GTCAAAATCTGGGAGGCCAATGAGTCCGAGACTTTTGTGACAGCTCGGATTTCCACC	480
Qy	481	AGAGAGAGGAGGACTTGGGCCAGTTTATGCTTCCAGTGGAGGCAATTTTGGGGCAGAA	540
Db	481	AGAGAGAGAGGAGGACTTGGGCCAGTTTATGCTTCCAGTGGAGGCAATTTTGGGGCAGAA	540
Qy	541	TACAGAGATATGGAATCAGATTATTCAGAGACAGGAGTTGACCAACTGCAAAAGATGATT	600
Db	541	TACAGAGATATGGAATCAGATTATTCAGAGACAGGAGTTGACCAACTGCAAAAGATGATT	600
Qy	601	GACACCATCAAAACCAACCTTGACGACAGAGAAATCATCATGTGCGTTGGAAATCCAGA	660
Db	601	GACACCATCAAAACCAACCTTGACGACAGAGAAATCATCATGTGCGTTGGAAATCCAGA	660
Qy	661	GATCTTCTCTATAGGCGCTGCTCCATAGCATGCGCCCTGCGGACTTATGTGGTAGAC	720
Db	661	GATCTTCTCTATAGGCGCTGCTCCATAGCATGCGCCCTGCGGACTTATGTGGTAGAC	720
Qy	721	AGTGAACCTGTCTGTGCAGCTGTACAGAGATCGGAGACATGGGAGCTCGGTGTGCTTTC	780
Db	721	AGTGAACCTGTCTGTGCAGCTGTACAGAGATCGGAGACATGGGAGCTCGGTGTGCTTTC	780
Qy	781	AACATGCGCAGGTACCGCTGTCTCAGTACATGATTTGGGCACATCAGGCGCTGAGGCA	840
Db	781	AACATGCGCAGGTACCGCTGTCTCAGTACATGATTTGGGCACATCAGGCGCTGAGGCA	840
Qy	841	GGTGACTTTATACACACTTTGGGAGATGACATATTATCTGAATCATGACGACCTG	900
Db	841	GGTGACTTTATACACACTTTGGGAGATGACATATTATCTGAATCATGACGACCTG	900
Qy	901	AAATTCACGCTTCAGCGAGAACCCAGCTTTTCCAAAGCTCAGGATTTCTTGAAGAGTT	960
Db	901	AAATTCACGCTTCAGCGAGAACCCAGCTTTTCCAAAGCTCAGGATTTCTTGAAGAGTT	960
Qy	961	GAGAAAAATTGATGACTTTAAAGCTGAAGACTTTCAGATTTGAAGGTTCAATCCGATCCA	1020
Db	961	GAGAAAAATTGATGACTTTAAAGCTGAAGACTTTCAGATTTGAAGGTTCAATCCGATCCA	1020
Qy	1021	ACATTTAAATATGGAAATGAGCTGTAGGGGCTTCAAAGAGGTGAGAGATATTGTCA	1080
Db	1021	ACATTTAAATATGGAAATGAGCTGTAGGGGCTTCAAAGAGGTGAGAGATATTGTCA	1080
Qy	1081	GTCCTTAAGGGGTGGGCTGAGATGCGAGAGTAAAGTTCTTTTGTCTTAAAGATAAAGG	1140
Db	1081	GTCCTTAAGGGGTGGGCTGAGATGCGAGAGTAAAGTTCTTTTGTCTTAAAGATAAAGG	1140
Qy	1141	AACATGATCAAAAAATCTTCCTGTGACCTATCATGATTATTAATTTTAAAGATGTGCACT	1200
Db	1141	AACATGATCAAAAAATCTTCCTGTGACCTATCATGATTATTAATTTTAAAGATGTGCACT	1200
Qy	1201	GAGCAATGTACTGTGCCAGTTCTTTTCATTAATTAAGGCTTTGATTTACTCAGTGAAG	1260
Db	1201	GAGCAATGTACTGTGCCAGTTCTTTTCATTAATTAAGGCTTTGATTTACTCAGTGAAG	1260
Qy	1261	GTAATCTGACATGCTGAGGTTATGAACAAAGTGAAGAGATGAATGTATGTCTTTAG	1320
Db	1261	GTAATCTGACATGCTGAGGTTATGAACAAAGTGAAGAGATGAATGTATGTCTTTAG	1320
Qy	1321	CAAAAAACATGTATGTGCATTTCAATCCGATCTTATTAAGAGGTTGTGTGAATTTCC	1380
Db	1321	CAAAAAACATGTATGTGCATTTCAATCCGATCTTATTAAGAGGTTGTGTGAATTTCC	1380
Qy	1381	AACCTATTTTGGAAATTTTATTAAGATTTTAAAGATTTTCAAGGATTTCCCTCAAT	1440
Db	1381	AACCTATTTTGGAAATTTTATTAAGATTTTAAAGATTTTCAAGGATTTCCCTCAAT	1440

CY	1441	CHAGAGGACCGAGTAAACCACTGCATCANGAGTGAGCGGTATGAACCTTA	1496
DB	1441	CTGAGGGAGCTAGATAACAACATCATCATGATGAGTTATGAACCTTA	1496
RESULT 4			
HSTSYN1		1536 bp	mRNA linear PRI 12-SEP-1993
LOCUS			
DEFINITION		Human mRNA for thymidylate synthase (EC 2.1.1.45).	
ACCESSION	X02308		
VERSION	X02308.1 GI:37478		
KEYWORDS	Inverted repeat; synthetase; tandem repeat.		
SOURCE	Homo sapiens.		
ORGANISM	Homo sapiens		
REFERENCE	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo. 1 (bases 1 to 1536) Takeishi,K., Kaneda,S., Ayusawa,D., Shimizu,K., Gotch,O. and Seno,T.		
AUTHORS			
TITLE	Nucleotide sequence of a functional cDNA for human thymidylate synthase		
JOURNAL	Nucleic Acids Res. 13 (6), 2035-2043 (1985)		
MEDLINE	85215597		
PUBMED	2987839		
COMMENT	Data kindly reviewed (22-OCT-1985) by Seno T.		
FEATURES			
source	Location/Qualifiers		
	1..1536	/organism="Homo sapiens"	
	/db_xref="taxon:9606"		
	14..103	/note="triple tandemly repeated elements"	
	misc_feature		
repeat_region	14..16	/note="direct repeat 1"	
	35..103	/note="stem-loop structure"	
	35..69	/note="pot. stem-loop structure"	
	35..41	/note="inverted repeat A"	
	42..44	/note="direct repeat 1"	
	63..103	/note="pot. stem-loop structure"	
	63..69	/note="inverted repeat A'"	
	70..72	/note="direct repeat 1"	
	97..103	/note="inverted repeat A'"	
	104..106	/note="direct repeat 1"	
	106..1047	/note="thymidylate synthase (aa 1-313)"	
	/codon_start=1		
	/protein_id="CAA26178.1"		
	/db_xref="GI:37479"		
	/db_xref="SWISS-PROT:P04818"		
	/translation="MPVAGSELPRRPDPAPAGRDAERPPHGELOYLGIOIHLKRC VRKDRVTGLTSVPGQARSLDERPILITTRFVPMGVABEILNTIKSTNAKEL SKGVYIDWANGSDFLDSIGFSREBGDLGPVGFMHPGAEYRMDSIDYSQGVDA LORVDITKNPDNRDITICAMNPREDIPALAPPCHALCOFYVNSELSCOLYRSGG KGLGPNPTNSAYALLTYMAHTITGAKRGDPIHTLGDPAHYLYNHLEPKIQLOGERPR PELRLLRKVERKIDDFEAEDFPQEGNEPHPTKKMAAV"		
	1231..1236	/note="pot. polyadenylation signal"	
	1519..1524	/note="pot. polyadenylation signal"	
	1536	/note="polyadenylation site"	
	369 c	399 g	378 t
BASE COUNT	390 a	399 g	378 t
ORIGIN			
Query Match		90.8%	Score 1394; DB 9; Length 1536;

AUTHORS Shichijo, S. and Itoh, K.
TITLE Identification of immuno-peptidomimics that recognized by
tumor-reactive CTL generated from TIL of colon cancer patients
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1567)
AUTHORS Itoh, K. and Shichijo, S.
TITLE Direct Submission
JOURNAL Submitted (24-May-2001) Shigeaki Shichijo, Kurume University School
of Medicine, Department of Immunology, 67 Asahi-machi, Kurume,
Fukuoka 830-0011, Japan (E-mail: shichijo@med.kurume-u.ac.jp,
Tel: 81-942-31-7551, Fax: 81-942-31-7699)
FEATURES
source
1. .1567
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="SW620"
/cell_type="colon adenocarcinoma"
gene
1. .1567
/gene="OK/SW-cl.29"
82. .1023
/gene="OK/SW-cl.29"
/codon_start=1
/product="thymidylate synthetase"
/protein_id="BAB93473.1"
/db_xref="GI:21104406"
/translation="MPVAGSELPRRLPPAORBDARPPHGLQYLQTOHILRG
VRKQDRTGTTISVFGMAQRYSLDERPLITTRVFWKSVLELLMFKSTAKELS
SKGVKIMDANGSRDPLDLPSTREBDLPVYGFVRMFRGAFRDMESYSGGVDO
LQVITDIKINPDDRIIMCAMPNDPPLMALPPCALQFYVNSLSGLQIRSD
MGLGVFNATSYALTYMIAHITGKPGDPIHTLGDPAHYNIHLPLKIQLOREPRP
PKRLIRKVEKIDDFKADPQIBGNPHPTIKEMAV"
BASE COUNT 431 a 366 c 389 g 381 t
ORIGIN
Query Match 77.4%; Score 1189; DB 9; Length 1567;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1339; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 37 CCGGCGCACTGGCTGCTCCGTCCGCGCGCGCACTTGGCTGCTCCGCTCCCGCC 96
DB 13 CCGGCGCACTGGCTGCTCCGTCCGCGCGCGCACTTGGCTGCTCCGCTCCCGCC 72
QY 97 CGCGCGCCATGCTGCTGCGCGCTGCGAGCTGCGCGCTTGGCTCCCGCGCGCA 156
DB 73 CGCGCGCCATGCTGCTGCGCGCTGCGAGCTGCGCGCTTGGCTCCCGCGCGCA 132
QY 157 CAGGAGCGGAGCGCGAGCGCGCTGCGCGCGCGCGAGCTGCGAGCTGCGAGCT 216
DB 133 CAGGAGCGGAGCGCGAGCGCGCTGCGCGCGCGCGAGCTGCGAGCTGCGAGCT 192
QY 217 CAACAATCTCTCGCTGCGCGCTGCGAGGAGGAGCGAGCGGAGCGGAGCGGAGCT 276
DB 193 CAACAATCTCTCGCTGCGCGCTGCGAGGAGGAGCGAGCGGAGCGGAGCGGAGCT 252
QY 277 GTATTGGGAGTGAAGCGCGCTCAAGCTGAGAGATTAATCCCTCTGCAACAACAA 336
DB 253 GTATTGGGAGTGAAGCGCGCTCAAGCTGAGAGATTAATCCCTCTGCAACAACAA 312
QY 337 CGTGTCTTGAAGGAGTGTGAGGAGTGTGCTGTGTTATCAAGGATTCACAAAT 396
DB 313 CGTGTCTTGAAGGAGTGTGAGGAGTGTGCTGTGTTATCAAGGATTCACAAAT 372
QY 397 GCTAAAGAGCTGCTTCAAGGAGTGAATCTGGGATGCCAATGGATTCCTT 456
DB 373 GCTAAAGAGCTGCTTCAAGGAGTGAATCTGGGATGCCAATGGATTCCTT 432
QY 457 TTGACAGCTTGGATTTCTCCACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 516
DB 433 TTGACAGCTTGGATTTCTCCACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 492
QY 517 CAGTGAAGGAGATTTGGGAGAGAGATTAAGAGATTAAGAGATTAAGAGATTAAGAG 576
DB 493 CAGTGAAGGAGATTTGGGAGAGAGATTAAGAGATTAAGAGATTAAGAGATTAAGAG 552

QY 577 GTTGACAGCTTGAAGAGATTTGACACCATCAAAACCTGACAGAGAGATTC 636
DB 553 GTTGACAGCTTGAAGAGATTTGACACCATCAAAACCTGACAGAGAGATTC 612
QY 637 ATCATGTGGCTTGAAGATTCAGAGATCTTCTGATGAGCGCTCCATGCCATGCC 696
DB 613 ATCATGTGGCTTGAAGATTCAGAGATCTTCTGATGAGCGCTCCATGCCATGCC 672
QY 697 CTCTGCAAGTCTTATGAGTGAACAGTGAAGCTGCTCCGACAGCTGTAACAGAGATCGGGA 756
DB 673 CTCTGCAAGTCTTATGAGTGAACAGTGAAGCTGCTCCGACAGCTGTAACAGAGATCGGGA 732
QY 757 GACATGGGCTGCTGAGTGCCTTTCAACATGCGCACTACGCTTCTGAGTACATGAT 816
DB 733 GACATGGGCTGCTGAGTGCCTTTCAACATGCGCACTACGCTTCTGAGTACATGAT 792
QY 817 GCGCACATCAAGGAGCTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 876
DB 793 GCGCACATCAAGGAGCTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 852
QY 877 TACCTGAATCATGAG 936
DB 853 TACCTGAATCATGAG 912
QY 937 AACCTGAGATCTTCCGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 996
DB 913 AACCTGAGATCTTCCGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 972
QY 997 ATTGAAGGATTAATCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1056
DB 973 ATTGAAGGATTAATCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1032
QY 1057 AAAGAGAGAGAGAGATTTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 1116
DB 1033 AAAGAGAGAGAGAGATTTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 1092
QY 1117 TCTTTTGTCTTAAAG 1176
DB 1093 TCTTTTGTCTTAAAG 1152
QY 1177 TTAATTTTAAAG 1236
DB 1153 TTAATTTTAAAG 1212
QY 1237 AGGCTTGAATTAATCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1296
DB 1213 AGGCTTGAATTAATCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1272
QY 1297 AGAATGAATGATAGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1356
DB 1273 AGAATGAATGATAGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1332
QY 1357 ATTAAGAGAGTGTGAGATTTTC 1378
DB 1333 ATTAAGAGAGTGTGAGATTTTC 1354
RESULT 9
AX069365 942 bp DNA linear PAT 25-JAN-2001
LOCUS AX069365
DEFINITION Sequence 29 from Patent WO0102600.
ACCESSION AX069365
VERSION AX069365.1 GI:12579227
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
AUTHORS Yuan, C.S.
TITLE Detection of analytes using attenuated enzymes
JOURNAL Patent: WO 0102600-A 29 11-JAN-2001;

QY 706 TTCTATGTGTGAACAGTGAAGCTGTCTCCAGCTGTATACAGAGATCGGAGACATYGGGC 765
 DB 601 TTCTATGTGTGAACAGTGAAGCTGTCTCCAGCTGTATACAGAGATCGGAGACATYGGGC 660
 QY 766 CTCGGTGTGCTCTTCAACATCGCAGCTACGCCCTGTCTCAAGTATGATTCGCGACATC 825
 DB 661 CTCGGTGTGCTCTTCAACATCGCAGCTACGCCCTGTCTCAAGTATGATTCGCGACATC 720
 QY 826 ACGGGCTGGAAGCCAGGTGACTTTATACACATTTGGAGATGACATATTTACTGAT 885
 DB 721 ACGGGCTGGAAGCCAGGTGACTTTATACACATTTGGAGATGACATATTTACTGAT 780
 QY 886 CACATCGAGCCAGTGAATAATTCAGTCTACGCGAGAACCCAGACTTTCCCAAGCTCGAG 945
 DB 781 CACATCGAGCCAGTGAATAATTCAGTCTACGCGAGAACCCAGACTTTCCCAAGCTCGAG 840
 QY 946 ATTCTTGGAAAAGTTGAGAAAATTTGATGACTTCAAGCTGAGACTTTTCAGATGAGGG 1005
 DB 841 ATTCTTGGAAAAGTTGAGAAAATTTGATGACTTCAAGCTGAGACTTTTCAGATGAGGG 900
 QY 1006 TACAATCCGATCCACACTTATTAATGAAATGGCTGTTAG 1047
 DB 901 TACAATCCGATCCACACTTATTAATGAAATGGCTGTTAG 942

 RESULT 11
 AX341901 AX341901 535 bp DNA linear PAT 10-JAN-2002
 LOCUS Sequence 2148 from Patent WO0196388.
 DEFINITION AX341901
 ACCESSION AX341901
 VERSION AX341901.1 GI:18137883
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1
 AUTHORS Jiang, Y., Harlocker, S.J. and Secrist, H.
 TITLE Compositions and methods for the therapy and diagnosis of colon
 cancer
 JOURNAL Patent: WO 0196388-A 2148 20-DEC-2001;
 CORIXA CORPORATION (US)
 FEATURES
 source 1..535
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /cell_line="taxon:9606" 118 t 2 others
 BASE COUNT 135 a 129 c 151 g 118 t
 ORIGIN
 Query Match 34.6%; Score 532; DB 6; Length 535;
 Best Local Similarity 100.0%; Pred. No. 1.9e-308; Indels 0; Gaps 0;
 Matches 532; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 203 ACCCTGGGGGAGATCCCAACATCTCCGCTGCGGCTGCGAGAGACACCGACGGGCA 262
 DB 1 ACCTGGGGGAGATCCCAACATCTCCGCTGCGGCTGCGAGAGACACCGACGGGCA 60
 QY 263 CCGGCAACCTGTGTGATTCGGGATGACGGCGCGCTACAGCTGAGAGATGATTCCTC 322
 DB 61 CCGGCAACCTGTGTGATTCGGGATGACGGCGCGCTACAGCTGAGAGATGATTCCTC 120
 QY 323 TGTGACACCAAAAGCTGTCTTGTGAAGGGTGTGTTGAGAGAGTGTGCTGTGTTTCA 382
 DB 121 TGTGACACCAAAAGCTGTCTTGTGAAGGGTGTGTTGAGAGAGTGTGCTGTGTTTCA 180
 QY 383 AGGATCCAGAAATGCTAAAGAGCTGTCTTCAAGGAGTGAATACTGGATGCAATG 442
 DB 181 AGGATCCAGAAATGCTAAAGAGCTGTCTTCAAGGAGTGAATACTGGATGCAATG 240
 QY 443 GATCCGAGACTTTTGGACAGCTGGAGATTCACACAGAGAAAGGGGACTTGGGCC 502
 DB 241 GATCCGAGACTTTTGGACAGCTGGAGATTCACACAGAGAAAGGGGACTTGGGCC 300

QY 503 CAGTTTANGGCTTCCAGTGGAGGCAATTTTGGGGCAGAAATACAGAGATGATGAAATCAGATT 562
 DB 301 CAGTTTANGGCTTCCAGTGGAGGCAATTTTGGGGCAGAAATACAGAGATGATGAAATCAGATT 360
 QY 563 ATTCAGAGCAGGGAGTTGACCAACTGCAAGAGTGAATGACACATCAAAACCACTTG 622
 DB 361 ATTCAGAGCAGGGAGTTGACCAACTGCAAGAGTGAATGACACATCAAAACCACTTG 420
 QY 623 ACCAGAGAAATCATCATGTCGCTTGGATTCAGAGATCTTCTGTATGCGCTGC 682
 DB 421 ACCAGAGAAATCATCATGTCGCTTGGATTCAGAGATCTTCTGTATGCGCTGC 480
 QY 683 CTCACATCCATCCCTCTGCGCAGTTCATATGTGTGAACAGTGAAGCTGTCTG 734
 DB 481 CTCACATCCATCCCTCTGCGCAGTTCATATGTGTGAACAGTGAAGCTGTCTG 532

 RESULT 12
 AB077207 693 bp mRNA linear PRI 09-JAN-2002
 LOCUS Homo sapiens TS mRNA for thymidylate synthase, complete cds,
 cell_line:MKM45.
 DEFINITION AB077207
 ACCESSION AB077207
 VERSION AB077207.1 GI:18150848
 KEYWORDS
 SOURCE Homo sapiens cell_line:MKM45 cDNA to mRNA.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1
 AUTHORS Hisatomi, H. and Nagaoka, K.
 TITLE Homo sapiens TS mRNA for thymidylate synthase
 JOURNAL Published Only in Database (2002)
 REFERENCE 2 (bases 1 to 693)
 AUTHORS Hisatomi, H. and Nagaoka, K.
 TITLE Direct Submission
 JOURNAL Submitted (07-JAN-2002) Hisashi Hisatomi, SRI, Inc., Center for
 Molecular Biology and CytoGenetics; 5-6-50 Shin-machi, Hino, Tokyo
 191-0002, Japan (E-mail: hisatomi@r1.sri-inc.co.jp,
 Tel:81426483793 (ex.7926), Fax:81426484043)
 FEATURES
 source 1..693
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="18"
 /map="18p11.32"
 /cell_line="MKM45"
 1..693
 /gene="TS"
 1..693
 /note="TS" alternative splicing variant The protein encoded
 by this gene is 83aa shorter than the full-length isoform.
 This variant uses an alternative splice site resulting in
 it lacking 249nt in the coding region compared to the
 full-length isoform reported in Acc# D00566"
 /codon_start=1
 /evidence=experimental
 /product="thymidylate synthase"
 /protein_id="BAB83676.1"
 /db_xref="GI:18150849"
 /translation="MPPASGSRPRPPPPAAGDAPRPPEHGLQYGOIHILRG
 VKMDRTGTGLSVFGQARYSLRYSQGVQDQKRVITDTIKNPDDREIIMCANMPR
 DPLALPCHALCOFYVNVSEISCOLYRSGMDGLVPENIASVALTYMAHTGL
 RKGDFHTIGDAHIYLNHLEPKIQLOREPRFPFKILIRKVKIDDFRAEDFOIEGY
 NHPITIMENAV"
 BASE COUNT 171 a 201 c 182 g 139 t
 ORIGIN
 Query Match 31.9%; Score 490; DB 9; Length 693;
 Best Local Similarity 100.0%; Pred. No. 4.2e-283; Indels 0; Gaps 0;
 Matches 490; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 558 AGATTATTCAGGACAGGAGTTGAACAATCGCAAAAGATGTAACAACATCAAAAACCA 617
 DB 204 AGATTATTCAGGACAGGAGTTGAACAATCGCAAAAGATGTAACAACATCAAAAACCA 263
 QY 618 CCCTGACGACAGAGAAATCATCATGTCGCTTGGAATCCAAAGATCTTCTCTGATGCG 677
 DB 264 CCCTGACGACAGAGAAATCATCATGTCGCTTGGAATCCAAAGATCTTCTCTGATGCG 323
 QY 678 GCTGCTCCATGACATGCGCTCTGCAATCTATGATGTAACAGTGAAGTCCGCA 737
 DB 324 GCTGCTCCATGACATGCGCTCTGCAATCTATGATGTAACAGTGAAGTCCGCA 383
 QY 738 GCTGACACAGAGATGAGGAGACAGGCGCTCGTGTGCTTCAACATCCAGTACG 797
 DB 384 GCTGACACAGAGATGAGGAGACAGGCGCTCGTGTGCTTCAACATCCAGTACG 443
 QY 798 CCTGCTCACTACATGATGTCGACATCAACGCGCTGAAAGCAGTGAATTAACAC 857
 DB 444 CCTGCTCACTACATGATGTCGACATCAACGCGCTGAAAGCAGTGAATTAACAC 503
 QY 858 TTGGAGATGACATATTTACTGTAATCAATCAAGCACTGAAATTCAGCTTGAGCG 917
 DB 504 TTGGAGATGACATATTTACTGTAATCAATCAAGCACTGAAATTCAGCTTGAGCG 563
 QY 918 AGAACCCAGACTTTCCCAAGCTCAGGATCTTGAAAGTTGAGAAATTTGATGACT 977
 DB 564 AGAACCCAGACTTTCCCAAGCTCAGGATCTTGAAAGTTGAGAAATTTGATGACT 623
 QY 978 CAAGCTGAAGCTTTCAGATGTAAGGATCAATCCGATCCAACTTAAATGGAAT 1037
 DB 624 CAAGCTGAAGCTTTCAGATGTAAGGATCAATCCGATCCAACTTAAATGGAAT 683
 QY 1038 GGCTGTTAG 1047
 DB 684 GGCTGTTAG 693

RESULT 13
 AR144965
 LOCUS AR144965 18596 bp DNA linear PAT 08-AUG-2001
 DEFINITION Sequence 11 from Patent US 6210950.
 ACCESSION AR144965
 VERSION AR144965.1 GI:15106832
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 18596)
 AUTHORS Johnson, W.G. and Stenroos, E.Scott.
 TITLE Methods for diagnosing, preventing, and treating developmental disorders due to a combination of genetic and environmental factors
 JOURNAL Patent: US 6210950-A 11 03-APR-2001;
 FEATURES
 source 1..18596
 /organism="Unknown"
 BASE COUNT 4521 a 3991 c 4479 g 5605 t
 ORIGIN

Query Match 31.7%; Score 487; DB 6; Length 18596;
 Best Local Similarity 99.7%; Pred. No. 4.5e-281;
 Matches 587; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 908 AGCTTCAGGAGAACCCAGACTTTCCCAAAGCTCAGATTTCTGAAAAGTTGAGAAA 967
 DB 15611 AGCTTCAGGAGAACCCAGACTTTCCCAAAGCTCAGATTTCTGAAAAGTTGAGAAA 15670
 QY 968 TTGATGACTTCAAAAGCTGAAGACTTTCAATTGAAGGATCAATCCGATCCACTATTA 1027
 DB 15671 TTGATGACTTCAAAAGCTGAAGACTTTCAATTGAAGGATCAATCCGATCCACTATTA 15730
 QY 1028 AAATGGAATGCTGTTAGGGTGTCTTCAAGAGCTGGAAGATATTTGCTGCTTTA 1087

DB 15731 AAATGGAATGCTGTTAGGGTGTCTTCAAGAGCTGGAAGATATTTGCTGCTTTA 15790
 QY 1088 GGGGTTGGGCTGATATCCGAGTAAAGTTCTTTTGTCTTAAAGAAAGAACTAGG 1147
 DB 15791 GGGGTTGGGCTGATATCCGAGTAAAGTTCTTTTGTCTTAAAGAAAGAACTAGG 15850
 QY 1148 TCMAAATCTGTCCGTAACTATCACTATTAATTTTAAAGATGTGCACTGGCAAT 1207
 DB 15851 TCMAAATCTGTCCGTAACTATCACTATTAATTTTAAAGATGTGCACTGGCAAT 15910
 QY 1208 GTAAGTGTCCAGTTCTTCCATATATAAAGCTTTAGTTAATCACTAGGATCTG 1267
 DB 15911 GTAAGTGTCCAGTTCTTCCATATATAAAGCTTTAGTTAATCACTAGGATCTG 15970
 QY 1268 ACAATGCTGAGGTTATGAACAAAGTGAAGAGATGAATGATGTGCTTTAGCAAAAAC 1327
 DB 15971 ACAATGCTGAGGTTATGAACAAAGTGAAGAGATGAATGATGTGCTTTAGCAAAAAC 16030
 QY 1328 ATGTATGTCAATTTCAATCCAGTACTATTAAGAAGAGTGTGATTTCAAGACTAT 1387
 DB 16031 ATGTATGTCAATTTCAATCCAGTACTATTAAGAAGAGTGTGATTTCAAGACTAT 16090
 QY 1388 TTTTGAATATTTTAAATATTTTAAAGATTTCAAGCTATTCCTCAATCTGAGGG 1447
 DB 16091 TTTTGAATATTTTAAATATTTTAAAGATTTCAAGCTATTCCTCAATCTGAGGG 16150
 QY 1448 AGCTGATGAACACCAATGAGATGATGATGATGATGATGATGATGATGATGAT 1496
 DB 16151 AGCTGATGAACACCAATGAGATGATGATGATGATGATGATGATGATGATGAT 16199

RESULT 14
 AX050451
 LOCUS AX050451 18596 bp DNA linear PAT 12-JAN-2001
 DEFINITION Sequence 11 from Patent WO0071754.
 ACCESSION AX050451
 VERSION AX050451.1 GI:12226657
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 18596)
 AUTHORS Johnson, W.G. and Stenroos, E.S.
 TITLE Methods for diagnosing, preventing, and treating developmental disorders due to a combination of genetic and environmental factors
 JOURNAL Patent: WO 0071754-A 11 30-NOV-2000;
 FEATURES
 source 1..18596
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 4521 a 3991 c 4479 g 5605 t
 ORIGIN

Query Match 31.7%; Score 487; DB 6; Length 18596;
 Best Local Similarity 99.7%; Pred. No. 4.5e-281;
 Matches 587; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 908 AGCTTCAGGAGAACCCAGACTTTCCCAAAGCTCAGATTTCTGAAAAGTTGAGAAA 967
 DB 15611 AGCTTCAGGAGAACCCAGACTTTCCCAAAGCTCAGATTTCTGAAAAGTTGAGAAA 15670
 QY 968 TTGATGACTTCAAAAGCTGAAGACTTTCAATTGAAGGATCAATCCGATCCACTATTA 1027
 DB 15671 TTGATGACTTCAAAAGCTGAAGACTTTCAATTGAAGGATCAATCCGATCCACTATTA 15730
 QY 1028 AAATGGAATGCTGTTAGGGTGTCTTCAAGAGCTGGAAGATATTTGCTGCTTTA 1087
 DB 15731 AAATGGAATGCTGTTAGGGTGTCTTCAAGAGCTGGAAGATATTTGCTGCTTTA 15790
 QY 1088 GGGGTTGGGCTGATATCCGAGTAAAGTTCTTTTGTCTTAAAGAAAGAACTAGG 1147

```

Db 15791 GGGGTTGGGCTGGATGCCGAGTTAAAGTTCCTTTTGGCTTAAAGAAAAGGACTAGG 15850
Qy 1148 TCAGAAATCTGTCCGTGACCTATCATGTTATATTTTAAAGATGTTGCCACTGCAAT 1207
Db 15851 TCAGAAATCTGTCCGTGACCTATCATGTTATATTTTAAAGATGTTGCCACTGCAAT 15910
Qy 1208 GTAACGTGCGAGTCTTTTCCATATATAAAGGCTTTGAGTTACTCACTGAGGGTATCTG 1267
Db 15911 GTAACGTGCGAGTCTTTTCCATATATAAAGGCTTTGAGTTACTCACTGAGGGTATCTG 15970
Qy 1268 ACAATGCTGAGGTTATGAACAAGTGAGGAAATGAATGATGATGCTCTTACCAAAAAC 1327
Db 15971 ACAATGCTGAGGTTATGAACAAGTGAGGAAATGAATGATGATGCTCTTACCAAAAAC 16030
Qy 1328 ATGTATGTCATTTTAAATCCACGCTACTTTTAAAGAGGTTGGGATTTTCAACAAGCTAT 1387
Db 16031 ATGTATGTCATTTTAAATCCACGCTACTTTTAAAGAGGTTGGGATTTTCAACAAGCTAT 16090
Qy 1388 TTTTGAATATTTTAAAGATTTTAAAGATTTTCAACAAGCTATTTCCCAATCTGAGGG 1447
Db 16091 TTTTGAATATTTTAAAGATTTTAAAGATTTTCAACAAGCTATTTCCCAATCTGAGGG 16150
Qy 1448 AGCTGAGTACACATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1496
Db 16151 AGCTGAGTACACATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 16199

RESULT 15
AX330682 18596 bp DNA linear PAT 09-JAN-2002
LOCUS AX330682
DEFINITION Sequence 1191 from Patent WO0194629.
ACCESSION AX330682
VERSION AX330682.1 GI:18103660
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
Horigan, S., Soppet, D.R. and Weaver, Z.
Cancer gene determination and therapeutic screening using signature
TITLE
Gene sets
JOURNAL Patent: WO 0194629-A, 1191, 13-DEC-2001;
Avalon Pharmaceuticals (US)
FEATURES
source 1..18596
location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 4521 a 3991 c 4479 g 5605 t
ORIGIN
Query Match 31.7%; Score 487; DB 6; Length 18596;
Best Local Similarity 99.7%; Pred. No. 4.5e-281;
Matches 587; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 908 AGCTTACGAGAACCCAGACCTTTCCAAAAGCTCAGATTTCTGAAAAGTTGAGAAA 967
Db 15611 AGCTTACGAGAACCCAGACCTTTCCAAAAGCTCAGATTTCTGAAAAGTTGAGAAA 15670
Qy 968 TTGATGACTTCAAGCTGAAGACTTTCAGATTGAAGGTAACATCCGATCCAACTATTA 1027
Db 15671 TTGATGACTTCAAGCTGAAGACTTTCAGATTGAAGGTAACATCCGATCCAACTATTA 15730
Qy 1028 AAATGGAATGCTGTTTAAAGGCTTTCAAGAGCTGGAAGATTTGTCAGTCTTTA 1087
Db 15731 AAATGGAATGCTGTTTAAAGGCTTTCAAGAGCTGGAAGATTTGTCAGTCTTTA 15790
Qy 1088 GGGGTTGGGCTGGATGCCGAGTTAAAGTCTTTTGGCTTAAAGAAAGAAAGAACTAGG 1147
Db 15791 GGGGTTGGGCTGGATGCCGAGTTAAAGTCTTTTGGCTTAAAGAAAGAAAGAACTAGG 15850
Qy 1148 TCAGAAATCTGTCCGTGACCTATCATGTTATATTTTAAAGATGTTGCCACTGCAAT 1207

```

```

Db 15851 TCAGAAATCTGTCCGTGACCTATCATGTTATATTTTAAAGATGTTGCCACTGCAAT 15910
Qy 1208 GTAACGTGCGAGTCTTTTCCATATATAAAGGCTTTGAGTTACTCACTGAGGGTATCTG 1267
Db 15911 GTAACGTGCGAGTCTTTTCCATATATAAAGGCTTTGAGTTACTCACTGAGGGTATCTG 15970
Qy 1268 ACAATGCTGAGGTTATGAACAAGTGAGGAAATGAATGATGATGCTCTTACCAAAAAC 1327
Db 15971 ACAATGCTGAGGTTATGAACAAGTGAGGAAATGAATGATGATGCTCTTACCAAAAAC 16030
Qy 1328 ATGTATGTCATTTTAAATCCACGCTACTTTTAAAGAGGTTGGGATTTTCAACAAGCTAT 1387
Db 16031 ATGTATGTCATTTTAAATCCACGCTACTTTTAAAGAGGTTGGGATTTTCAACAAGCTAT 16090
Qy 1388 TTTTGAATATTTTAAAGATTTTAAAGATTTTCAACAAGCTATTTCCCAATCTGAGGG 1447
Db 16091 TTTTGAATATTTTAAAGATTTTAAAGATTTTCAACAAGCTATTTCCCAATCTGAGGG 16150
Qy 1448 AGCTGAGTACACATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1496
Db 16151 AGCTGAGTACACATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 16199

Search completed: November 29, 2002, 02:29:06
UDB time : 3980.66 secs

```

DR P-PSDB; AAU87071.
 XX New nucleic acid molecule Hong Kong New Gene 1 (HKNG1), useful for
 PT screening for molecules which modulate HKNG1 expression for the
 PT treatment of bipolar disorder and schizophrenia -
 XX Disclosure; Fig 45; 367pp; English.
 XX The invention relates to an isolated nucleic acid molecule comprising a
 CC nucleotide sequence that encodes a Hong Kong New Gene (HKNG) 1 gene
 CC product. The human gene for HKNG1 is located on chromosome 18p in
 CC an area associated with bipolar affective disorder, BAD. Also
 CC included are an expression vector comprising the nucleic acid, a
 CC host cell expressing the nucleic acid, an anti-HKNG1 antibody, a method
 CC of identifying modulators of HKNG1, and identifying an individual at
 CC risk of having HKNG1-mediated disorder comprising detecting the presence
 CC or absence of a polymorphism that correlates with an HKNG1 allele
 CC associated with the disorder, where the presence of the polymorphism
 CC indicates that the individual is at risk of having HKNG1-mediated
 CC disorder. A (small molecule) compound which modulates (inhibits or
 CC potentiates) expression of a HKNG1 gene or gene product in a human
 CC individual is useful for the treatment of a HKNG1-mediated disorder,
 CC such as bipolar affective disorder (BAD), severe bipolar affective
 CC disorder (BP-I) and schizophrenia. The present sequence is the cDNA
 CC encoding thymidylate synthase, TS. The gene for TS
 CC overlaps that of HKNG1 and therefore TS may also be involved in the
 CC diseases listed above.
 XX
 SQ Sequence 1536 BP; 390 A; 369 C; 399 G; 378 T; 0 other;
 Query Match 90.8%; Score 1394; DB 24; Length 1536;
 Best Local Similarity 99.9%; Pred.No. 0;
 Matches 1494; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 GGGGGGGGGGGGACCACTGGCCGCGCTCCGCGCCGCGCACTTGGGCTGCTCGCT 60
 Db 1 GGGGGGGGGGGGACCACTGGCTGCTCGCTCCGCGCGCGCACTTGGGCTGCTCGCT 60
 QY 61 CCGCGCGCGCACTTGGCTGCTCGCTCCGCGCGCGCGCGCGCGCATGCTGCGCGCGC 120
 Db 61 CCGCGCGCGCACTTGGCTGCTCGCTCCGCGCGCGCGCGCGCGCGCATGCTGCGCGCGC 120
 QY 121 TCGAGCTGCT 180
 Db 121 TCGAGCTGCT 180
 QY 181 CCGCT 240
 Db 181 CCGCT 240
 QY 241 AGAAGAGAGACCGGACCGGAGACCGGAGACCGGAGATTCGGGATTCGGGAGCGGCGCTAC 300
 Db 241 AGAAGAGAGACCGGACCGGAGACCGGAGACCGGAGATTCGGGATTCGGGAGCGGCGCTAC 300
 QY 301 AGCTGAGAGAGATTCCTGCTGCTGCAACAACAGGATTCGTTGTTGTTGTTGTTGTTG 360
 Db 301 AGCTGAGAGAGATTCCTGCTGCTGCAACAACAGGATTCGTTGTTGTTGTTGTTGTTG 360
 QY 361 GAGGAGTTGCTGCTGTTTATCAAGAGATCAACAATGCTAAAGAGCTGCTTCCAGAGGA 420
 Db 361 GAGGAGTTGCTGCTGTTTATCAAGAGATCAACAATGCTAAAGAGCTGCTTCCAGAGGA 420
 QY 421 GTGAATATCTGGATGCCAATGATCCCGAATTGTTGACAGCTTGGGATTCCTCACCC 480
 Db 421 GTGAATATCTGGATGCCAATGATCCCGAATTGTTGACAGCTTGGGATTCCTCACCC 480
 QY 481 AGAAGAGAGAGGAGCTTGGGCGCACTTATGCTTCCAGTGAAGCACTTTGGGCGAGAA 540
 Db 481 AGAAGAGAGAGGAGCTTGGGCGCACTTATGCTTCCAGTGAAGCACTTTGGGCGAGAA 540
 QY 541 TACAGAGATATGAGATCAATATATCAAGACAGGAGATTCGCAACATGCAAGAGATGAT 600
 Db 541 TACAGAGATATGAGATCAATATATCAAGACAGGAGATTCGCAACATGCAAGAGATGAT 600

QY 601 GACACCATCAAAACCAACCTGACGACAGAGAAATCATATGTCGCTTGGATCCAGAA 660
 Db 601 GACACCATCAAAACCAACCTGACGACAGAGAAATCATATGTCGCTTGGATCCAGAA 660
 QY 661 GATCTTCCCTGATGAGGCTGCTGCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 Db 661 GATCTTCCCTGATGAGGCTGCTGCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 QY 721 AGTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
 Db 721 AGTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
 QY 781 AACATGCGCAGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
 Db 781 AACATGCGCAGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
 QY 841 GGTGACTTATACACACTTGGGAGATGACATATTTACCTGATACATGACGACACTG 900
 Db 841 GGTGACTTATACACACTTGGGAGATGACATATTTACCTGATACATGACGACACTG 900
 QY 901 AAATTCAGCTTACGCGAGAAACCGAAGCTTCCAAAGCTCAGATTTCTCGAAAGTT 960
 Db 901 AAATTCAGCTTACGCGAGAAACCGAAGCTTCCAAAGCTCAGATTTCTCGAAAGTT 960
 QY 961 GAGAAATGATGACTTCAAGCTGAGACTTTCAGATTGAGGATGCAATCCGATCCA 1020
 Db 961 GAGAAATGATGACTTCAAGCTGAGACTTTCAGATTGAGGATGCAATCCGATCCA 1020
 QY 1021 ACTATTTAAATGAAATGCTGTTTAAAGGCTGTTTCAAAAGAGCTGAGAGATTTGTCA 1080
 Db 1021 ACTATTTAAATGAAATGCTGTTTAAAGGCTGTTTCAAAAGAGCTGAGAGATTTGTCA 1080
 QY 1081 GTCTTTAGGGGTTGGGCTGATGCGAGGTAAGTTCTTTTCTGCTTAAAGAAAGG 1140
 Db 1081 GTCTTTAGGGGTTGGGCTGATGCGAGGTAAGTTCTTTTCTGCTTAAAGAAAGG 1140
 QY 1141 AACTAGGTCAAAATCTGCTCGTGAACCTATGATTAATTTTAAAGAGTTGCGACCT 1200
 Db 1141 AACTAGGTCAAAATCTGCTCGTGAACCTATGATTAATTTTAAAGAGTTGCGACCT 1200
 QY 1201 GCGAAATGTAAGTGCAGATTCCTTCCATTAATAAAGGCTTGAAGTAACTCACTGAGG 1260
 Db 1201 GCGAAATGTAAGTGCAGATTCCTTCCATTAATAAAGGCTTGAAGTAACTCACTGAGG 1260
 QY 1261 GTATCGAATGCTGAGTATTAAGCAAGTGAAGATGAATGATGCTGCTG 1320
 Db 1261 GTATCGAATGCTGAGTATTAAGCAAGTGAAGATGAATGATGCTGCTG 1320
 QY 1321 CAAAACATGATGTCATTCATCCACGTAATTAAAGAGTTGTTGATTTTCAAC 1380
 Db 1321 CAAAACATGATGTCATTCATCCACGTAATTAAAGAGTTGTTGATTTTCAAC 1380
 QY 1381 AAGCTATTTTGGAAATTTTAAATTTTAAATTTTAAATTTTCAAGCTATTCCTCTCAAT 1440
 Db 1381 AAGCTATTTTGGAAATTTTAAATTTTAAATTTTAAATTTTCAAGCTATTCCTCTCAAT 1440
 QY 1441 CTGAGGAGCTGATGATCAATGATCATGATGATGATGATGATGATGATGATGATGAT 1496
 Db 1441 CTGAGGAGCTGATGATCAATGATCATGATGATGATGATGATGATGATGATGATGAT 1496
 RESULT 2
 AAS94945
 ID AAS94945 standard; DNA; 3298 BP.
 XX AAS94945;
 AC AAS94945;
 XX 14-FEB-2002 (first entry)
 DT Human DNA sequence #200 expressed during foam cell differentiation.
 XX Human; foam cell differentiation; atherosclerosis; cerebral stroke;
 KM

RESULT 3
AA84960/c
ID AA84960 standard; cDNA; 1539 BP.
XX
AC AA84960;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #20764.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW Food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYTE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI: 2001-639362/73.
DR P-Psdb; ABG20773.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 1; SEQ ID No 20764; 103bp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AA8497-AA84954 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1539 BP; 396 A; 379 C; 347 G; 417 T; 0 other;
XX
Query Match 34.9%; Score 536; DB 23; Length 1539;
Best Local Similarity 100.0%; Pred. No. 2e-266;
Matches 536; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 226 CTCGCTGCGGCTTCAGAGAGACGACCGGACCGGACCGGACCGGATTTGGC 285
D 1419 CTCGCTGCGGCTTCAGAGAGAGACGACCGGACCGGACCGGATTTGGC 1360
QY 286 ATGCAGGCGGCTTCAGAGAGAGATTCCTGCTGCAACAACGATGTTTC 345
D 1359 ATGCAGGCGGCTTCAGAGAGAGATTCCTGCTGCAACAACGATGTTTC 1300
QY 346 TGAAGGCTGTTTGGAGAGATTCCTGCTGCAACAACGATGTTTC 405
D 1299 TGAAGGCTGTTTGGAGAGATTCCTGCTGCAACAACGATGTTTC 1240
QY 406 CTGCTTCCAAAGGAGTGAATCTGGAATGCAATGCAATCTTTGACAGC 465
D 1239 CTGCTTCCAAAGGAGTGAATCTGGAATGCAATGCAATCTTTGACAGC 1180
QY 466 CTGGAATTCCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 525
D 1179 CTGGAATTCCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1120
QY 526 CATTTTGGGCGAGAGATTCAGAGATTCAGAGATTCAGAGATTCAGAG 585
D 1119 CATTTTGGGCGAGAGATTCAGAGATTCAGAGATTCAGAGATTCAGAG 1060
QY 586 CTGCAAGAGTGAATTCAGAGATTCAGAGATTCAGAGATTCAGAGATTC 641
D 1059 CTGCAAGAGTGAATTCAGAGATTCAGAGATTCAGAGATTCAGAGATTC 1004
RESULT 4
ABL38559
ID ABL38559 standard; cDNA; 535 BP.
XX
AC ABL38559;
XX
DT 08-APR-2002 (first entry)
XX
D Human colon tumour antigen polynucleotide SEQ ID NO:2148.
XX
KW Human; colon cancer; colon tumour antigen; cytostatic; vaccine;
KW colon tumour metastatic antigen; diagnosis; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200196388-A2.
XX
PD 20-DEC-2001.
XX
PF 08-JUN-2001; 2001WO-US18557.
XX
PR 09-JUN-2000; 2000US-210899P.
PR 20-FEB-2001; 2001US-270216P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Jiang Y, Harlocker SL, Secrist H;
XX
DR WPI: 2002-114514/15.
XX
PT Novel isolated colon tumor polynucleotide differentially expressed in
PT colon tumor or colon metastatic tumor and polypeptides encoded by them,
PT useful for inhibiting development of cancer in patient -
XX
PS Claim 1; SEQ ID 2148; 105bp; English.
XX
CC ABL38642 to ABL38645 represent human colon tumour antigen cDNA clones (I)
CC which were isolated from human colon tumour and colon metastatic tumour
CC cDNA libraries. (I) have cytostatic activity and can be used in vaccine
CC production. (I) can be used for stimulating and/or expanding T cells
CC specific for a tumour protein on contact with the T cells. They are also
CC useful for inhibiting the development of cancer in a patient. (I) can be
CC used as probes or primers for nucleic acid hybridisation, for preparing

CC mutant species primers, or primers for use in genetic constructions. (I)
CC can be used in the diagnosis of a colon tumor.

XX Sequence 535 BP; 135 A; 129 C; 151 G; 118 T; 2 other;

Query Match 34.6%; Score 532; DB 24; Length 535;

Best Local Similarity 100.0%; Pred. No. 2.3e-244;
Matches 532; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 203 ACCCTGAGGAGATCCAAACATCTCTCCGCTGCGGGCTCAGGAAGACACCGCACGGCA 262
DB 1 ACCTGGGGGAGATCCAAACATCTCTCCGCTGCGGGCTCAGGAAGACACCGCACGGCA 60
QY 263 CCGGACACCTGTGGTATTCCGCAATGAGCGCGCTACACGCTTGAGATTAATTCCTC 322
DB 61 CCGGACACCTGTGGTATTCCGCAATGAGCGCGCTACACGCTTGAGATTAATTCCTC 120
QY 323 TGTGACACCAACAGTGTCTGGAAGGCTTTTGAAGAGATGCTGTGATTATCA 382
DB 121 TGTGACACCAACAGTGTCTGGAAGGCTTTTGAAGAGATGCTGTGATTATCA 180
QY 383 AGGGATCCAAATGCTAAAGCTGTCTTCCAAGGAGTGAATTCGGATGCCAATG 442
DB 181 AGGGATCCAAATGCTAAAGCTGTCTTCCAAGGAGTGAATTCGGATGCCAATG 240
QY 443 GATCCGAGACTTTTGAACAGCTGGGATTTCTCAACAGAGAGAGGAGACTTGGGCC 502
DB 241 GATCCGAGACTTTTGAACAGCTGGGATTTCTCAACAGAGAGAGGAGACTTGGGCC 300
QY 503 CAGTTATGCTTCCAGTGGAGGCAATTTGGGGGAGAAATACAGAGATATGGAATCAGATT 562
DB 301 CAGTTATGCTTCCAGTGGAGGCAATTTGGGGGAGAAATACAGAGATATGGAATCAGATT 360
QY 563 ATTGAGAGACAGGAGATTGACCAACTGCAAGAGATGATTGACACCATCAAAACCAACCTG 622
DB 361 ATTGAGAGACAGGAGATTGACCAACTGCAAGAGATGATTGACACCATCAAAACCAACCTG 420
QY 623 ACGACAGAGATATCATATGTCGCTTGGAAATCCAGAGATCTTCTCGATGGCGCTGC 682
DB 421 ACGACAGAGATATCATATGTCGCTTGGAAATCCAGAGATCTTCTCGATGGCGCTGC 480
QY 683 CTCATGCCATGCGCTCTGCGAGTCTATGAGTGAACAGTAGAGCTGCTCG 734
DB 481 CTCATGCCATGCGCTCTGCGAGTCTATGAGTGAACAGTAGAGCTGCTCG 532

RESULT 5
AAF31109
ID AAF31109 standard; cDNA; 18596 BP.
XX
AC AAF31109;
XX
DT 27-APR-2001 (first entry)
XX
DE Thymidylate synthase coding sequence.
XX
KW Analyte-binding enzyme; analyte analysis; ss.
XX
OS Homo sapiens.
XX
PN W02001.02600-A2.
XX
PD 11-JAN-2001.
XX
PF 30-JUN-2000; 2000WO-US18057.
XX
PR 06-JUN-1999; 99US-0347878.
XX
PR 06-DEC-1999; 99US-0457205.
XX
PA (GENAT) GEN ATOMICS.
XX
PI Yuan C;
XX
```

DR WPI: 2001-071583/08.
XX Assaying method, useful for prognosis and diagnosis of disease,
PT comprises contacting sample with a mutant analyte-binding enzyme and
PT detecting binding -
XX
PS Disclosure; Page -: 187pp; English.
XX

CC The present invention relates to a method for assaying an analyte in a
CC sample comprising: contacting the sample with a mutant analyte-binding
CC enzyme which has binding affinity for the analyte or an immediate
CC analyte enzymatic conversion product but has attenuated catalytic
CC activity; and detecting resulting binding. The method is useful in
CC monitoring biological systems/processes, or prognosis/diagnosis of
CC disease caused by imbalances of the analytes. The present sequence is
CC a coding sequence used in the present invention.
CC Note: the present sequence is not shown in the specification, but was
CC from Genbank, using information given in the specification.

XX Sequence 18596 BP; 4521 A; 3991 C; 4479 G; 5605 T; 0 other;

Query Match 31.7%; Score 487; DB 22; Length 18596;
Best Local Similarity 99.7%; Pred. No. 4.6e-241;
Matches 587; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 908 AGCTTAGCGAGAACCCAGACCTTCCCAAGCTCAGAGATTTCTTGAAAGTGAAGAAA 967
DB 15611 AGCTTAGCGAGAACCCAGACCTTCCCAAGCTCAGAGATTTCTTGAAAGTGAAGAAA 15670
QY 968 TTGATGACTTCAAGCTGAAGACTTTGAGTTGAGGAGTCAATCCGACCACTATTA 1027
DB 15671 TTGATGACTTCAAGCTGAAGACTTTGAGTTGAGGAGTCAATCCGACCACTATTA 15730
QY 1028 AAATGSAATAGGCTGTTTGGGCTTCAAGGAGCTGAGAGATATGTCAGCTTTA 1087
DB 15731 AAATGSAATAGGCTGTTTGGGCTTCAAGGAGCTGAGAGATATGTCAGCTTTA 15790
QY 1088 GGGGTGGGCTGGAGTCCGAGGTAAAGTCTTTTCTCTTAAAGAAAGAACTAGG 1147
DB 15791 GGGGTGGGCTGGAGTCCGAGGTAAAGTCTTTTCTCTTAAAGAAAGAACTAGG 15850
QY 1148 TCAAAAATCTGTCGAGACCTATCAGTTAATTTTAAAGATGTTGCACTGSCAAAT 1207
DB 15851 TCAAAAATCTGTCGAGACCTATCAGTTAATTTTAAAGATGTTGCACTGSCAAAT 15910
QY 1208 GTAATGTCGAGTTCTTTCCATTAATTAAGAGCTTTGATTAATCTGAGAGGATCTG 1267
DB 15911 GTAATGTCGAGTTCTTTCCATTAATTAAGAGCTTTGATTAATCTGAGAGGATCTG 15970
QY 1268 ACAATGCTGAGTTATGAACAAGTGAAGAGATGAATGATGCTTGAACAAAC 1327
DB 15971 ACAATGCTGAGTTATGAACAAGTGAAGAGATGAATGATGCTTGAACAAAC 16030
QY 1328 ATGTATGCAATTCATCCACGACTTATTAAGAGAGTGGTGAATTCACAAGTAT 1387
DB 16031 ATGTATGCAATTCATCCACGACTTATTAAGAGAGTGGTGAATTCACAAGTAT 16090
QY 1388 TTTTGAATATTTTGAATATTTTGAATTTTGAATTTTGAATTTTGAATTTTGA 1447
DB 16091 TTTTGAATATTTTGAATATTTTGAATTTTGAATTTTGAATTTTGAATTTTGA 16150
QY 1448 AGCTGAGTAACACCATGATCATGATGATGATGATGATGATGATGATGATGAT 1496
DB 16151 AGCTGAGTAACACCATGATCATGATGATGATGATGATGATGATGATGATGAT 16199

RESULT 6
AAC91215
ID AAC91215 standard; DNA; 18596 BP.
XX
AC AAC91215;
XX
DT 20-MAR-2001 (first entry)
XX
```


XX DE Human thymidylate synthase gene SEQ ID NO: 11.
 XX KW Human; schizophrenia; developmental disorder; spina bifida cystica;
 KW Tourette's syndrome; bipolar illness; autism; conduct disorder;
 KW attention deficit disorder; obsessive compulsive disorder;
 KW chronic multiple tic syndrome; learning disorder; polymorphism; ds.
 XX OS Homo sapiens.
 XX PN WO200071754-A1.
 XX PD 30-NOV-2000.
 XX PF 24-MAY-2000; 2000WO-US14354.
 XX PR 25-MAY-1999; 99US-0318448.
 XX PA (UYME-) UNIV NEW JERSEY MEDICINE & DENTISTRY.
 PI Johnson WG, Steenroos ES;
 PI MPI; 2001-025174/03.
 DR PT Diagnosing a developmental disorder, e.g. schizophrenia, by forming
 XX PT datasets (DS) of genetic (e.g. genotypes of folate metabolism alleles)
 XX PT and environmental variables affecting an individual and then comparing
 XX PT these DS with reference DS -
 XX PS Disclosure; Page 125-131, 156pp; English.
 CC The present invention provides a novel method of estimating the
 CC susceptibility of an individual to a developmental disorder using genetic
 CC and environmental variables. The method can be used in the diagnosis,
 CC prevention and treatment of disorders such as schizophrenia, spina bifida
 CC cystica, Tourette's syndrome, bipolar illness, autism, conduct disorders,
 CC attention deficit disorder, obsessive compulsive disorder, chronic
 CC multiple tic syndrome and learning disorders such as dyslexia.
 XX SQ Sequence 18596 BP; 4521 A; 3991 C; 4479 G; 5605 T; 0 other;
 Query Match 31.7%; Score 487; DB 22; Length 18596;
 Best Local Similarity 99.7%; Pred. No. 4,6e-241;
 Matches 587; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 908 AGCTTCACCGAAGACCCGACCTTTCCCAAGCTCAGATTCTTCGAAAAGTTGAGAAA 967
 DB 15611 AGCTTCACCGAAGACCCGACCTTTCCCAAGCTCAGATTCTTCGAAAAGTTGAGAAA 15670
 QY 968 TTGATGATTCGAAGCTGAAGACTTCAGATTGAAGGGGTACAAATCCGATCCACTATTA 1027
 DB 15671 TTGATGATTCGAAGCTGAAGACTTCAGATTGAAGGGGTACAAATCCGATCCACTATTA 15730
 QY 1028 AATGGAATGGCTGTTTAAAGGCTTTAAAGAGCTGGAAGATATTGTCACTTTTA 1087
 DB 15731 AATGGAATGGCTGTTTAAAGGCTTTAAAGAGCTGGAAGATATTGTCACTTTTA 15790
 QY 1088 GGGGTTGGGCTGATGCGGAGTAAAGTTCTTTTGTCTTAAAGAAAGAAAGAACTAGG 1147
 DB 15791 GGGGTTGGGCTGATGCGGAGTAAAGTTCTTTTGTCTTAAAGAAAGAAAGAACTAGG 15850
 QY 1148 TCAAAATCTGTCCGTGACTATCACTAATTAATTTTAAAGATGTTGCCACTGCAAT 1207
 DB 15851 TCAAAATCTGTCCGTGACTATCACTAATTAATTTTAAAGATGTTGCCACTGCAAT 15910
 QY 1208 GTAATCTGTCCGTGACTATCACTAATTAATTTTAAAGATGTTGCCACTGCAAT 1267
 DB 15911 GTAATCTGTCCGTGACTATCACTAATTAATTTTAAAGATGTTGCCACTGCAAT 15970
 QY 1268 ACAAATCGAGGTTATGAAGAAAGTGAAGAAAGAAATGATGCTCTTACCAAAAAC 1327
 DB 15971 ACAAATCGAGGTTATGAAGAAAGTGAAGAAAGAAATGATGCTCTTACCAAAAAC 16030

QY 1328 ATGTATGTCATTTCAATCCGACTTATTAAGAAGTTGGTGAATTTCAACGACTAT 1387
 DB 16031 ATGTATGTCATTTCAATCCGACTTATTAAGAAGTTGGTGAATTTCAACGACTAT 16090
 QY 1388 TTTTGGAAATTTTGAATTTTAAAGATTTCAAGGTTATTCCTCAATCTGAGGG 1447
 DB 16091 TTTTGGAAATTTTGAATTTTAAAGATTTTAAAGATTTTCAAGGTTATTCCTCAATCTGAGGG 16150
 QY 1448 AGCTGATTAACCAATCGATCATGATGATGATGATGATGATGATGATGATGATGAT 1496
 DB 16151 AGCTGATTAACCAATCGATCATGATGATGATGATGATGATGATGATGATGATGAT 16199
 RESULT 7
 ID ABN95092 standard; DNA; 18596 BP.
 XX AC ABN95092;
 XX DT 13-AUG-2002 (first entry)
 XX DE Gene #1590 used to diagnose liver cancer.
 XX KW Gene; liver cancer; ds; hepatocellular carcinoma; hepatocytic;
 KW metastatic liver tumour; cytostatic; expression profile; disease state;
 KW disease progression; drug toxicity; drug efficacy; drug metabolism.
 XX OS Homo sapiens.
 XX PN WO200229103-A2.
 XX PD 11-APR-2002.
 XX PF 02-OCT-2001; 2001WO-US30589.
 XX PR 02-OCT-2000; 2000US-237054P.
 XX PA (GENE-) GENE LOGIC INC.
 XX PI Horne D, Alvares C, Peres-Da-Silva S, Vockley TG;
 XX WP1; 2002-426119/45.
 DR PT Diagnosing and detecting the progression of liver cancer,
 PT hepatocellular carcinoma or metastatic liver tumor in a patient,
 PT involves detecting the level of expression of two or more genes in a
 PT liver tissue sample -
 XX PS Claim 1; SEQ ID NO 1590; 298bp; English.
 CC The invention relates to a novel method for diagnosing and detecting the
 CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
 CC tumour in a patient, and differentiating metastatic liver cancer from
 CC hepatocellular carcinoma in a patient, involving detecting the level of
 CC expression of two or more genes represented in ABN9503-ABN97455 in a
 CC tissue sample. The method of the invention has hepatocytic, and
 CC cytotretic activity. The method is useful for diagnosing and detecting
 CC the progression of liver cancer, hepatocellular carcinoma and metastatic
 CC liver carcinoma in a patient. The method is useful for identifying
 CC expression profiles which serve as useful diagnostic markers as well as
 CC markers that can be used to monitor disease states, disease progression,
 CC drug toxicity, drug efficacy and drug metabolism.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIP0
 CC at ftp.wipo.int/pub/published_pat_sequences.
 SQ Sequence 18596 BP; 4521 A; 3991 C; 4479 G; 5605 T; 0 other;
 Query Match 31.7%; Score 487; DB 24; Length 18596;
 Best Local Similarity 99.7%; Pred. No. 4,6e-241;
 Matches 587; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 908 AGCTTCACCGAAGACCCGACCTTTCCCAAGCTCAGATTCTTCGAAAAGTTGAGAAA 967

RESULT 9
 ABL62854
 ID ABL62854 standard; DNA; 18596 BP.
 AC ABL62854;
 XX
 DT 15-MAY-2002 (first entry)
 XX
 DE Breast cancer related gene sequence SEQ ID NO:1191.
 XX
 KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
 KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
 KW cytostatic; gene therapy; antineoplastic; Wilms' tumour; adenocarcinoma;
 KW gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN WC200194629-A2.
 XX
 PD 13-DEC-2001.
 XX
 PF 30-MAY-2001; 2001WO-US10838.
 XX
 PR 05-JUN-2000; 2000US-209473P.
 PR 05-JUN-2000; 2000US-209531P.
 PR 18-SEP-2000; 2000US-233133P.
 PR 18-SEP-2000; 2000US-233617P.
 PR 20-SEP-2000; 2000US-234009P.
 PR 20-SEP-2000; 2000US-234034P.
 PR 20-SEP-2000; 2000US-234052P.
 PR 22-SEP-2000; 2000US-234509P.
 PR 22-SEP-2000; 2000US-234567P.
 PR 25-SEP-2000; 2000US-234923P.
 PR 25-SEP-2000; 2000US-234924P.
 PR 25-SEP-2000; 2000US-235077P.
 PR 25-SEP-2000; 2000US-235082P.
 PR 25-SEP-2000; 2000US-235134P.
 PR 25-SEP-2000; 2000US-235280P.
 PR 26-SEP-2000; 2000US-235637P.
 PR 27-SEP-2000; 2000US-235638P.
 PR 27-SEP-2000; 2000US-235711P.
 PR 27-SEP-2000; 2000US-235720P.
 PR 27-SEP-2000; 2000US-235840P.
 PR 27-SEP-2000; 2000US-235863P.
 PR 28-SEP-2000; 2000US-236028P.
 PR 28-SEP-2000; 2000US-236032P.
 PR 28-SEP-2000; 2000US-236033P.
 PR 28-SEP-2000; 2000US-236034P.
 PR 28-SEP-2000; 2000US-236109P.
 PR 28-SEP-2000; 2000US-236111P.
 PR 29-SEP-2000; 2000US-236842P.
 PR 29-SEP-2000; 2000US-236843P.
 PR 29-SEP-2000; 2000US-236891P.
 PR 02-OCT-2000; 2000US-237172P.
 PR 02-OCT-2000; 2000US-237173P.
 PR 02-OCT-2000; 2000US-237278P.
 PR 02-OCT-2000; 2000US-237294P.
 PR 02-OCT-2000; 2000US-237295P.
 PR 02-OCT-2000; 2000US-237316P.
 PR 03-OCT-2000; 2000US-237425P.
 PR 03-OCT-2000; 2000US-237598P.
 PR 03-OCT-2000; 2000US-237604P.
 PR 03-OCT-2000; 2000US-237606P.
 PR 03-OCT-2000; 2000US-237608P.
 PR 01-NOV-2000; 2000US-244867P.
 PR 01-NOV-2000; 2000US-245084P.
 XX
 PA (AVALON PHARM.
 XX
 PI Young PJ, Augustus M, Carter KC, Ebnner R, Endress G, Horrigan S,
 PI Sopet DR, Weaver Z;
 XX
 DR WPI; 2002-188264/24.
 XX

PR Screening for anti-neoplastic agent involves exposing cells to a
 PR chemical agent to be tested for anti-neoplastic activity, and
 PR determining a change in expression of a gene of a signature gene set
 XX
 PS Claim 1; SEQ ID 1191; 44pp; English.

CC The present invention describes a method (M1) for screening for an
 CC anti-neoplastic agent. The method involves exposing cells to a chemical
 CC agent to be tested for anti-neoplastic activity, determining a change in
 CC expression of at least one gene (I) of a signature gene set, where (I)
 CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
 CC to ABL70110), or is at least 95% identical to (S), where a change in
 CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
 CC activity and can be used in gene therapy. M1 can be used for screening
 CC an anti-neoplastic agent, and can be used for producing a product which
 CC is the data collected with respect to the anti-neoplastic agent as a
 CC result of M1, and the data is sufficient to convey the chemical
 CC structure and/or properties of the agent. M1 can be used in the
 CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
 CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
 CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
 CC carcinoma, papillary carcinoma and Wilms' tumour.

CC Sequence 18596 BP; 4521 A; 3991 C; 4479 G; 5605 T; 0 other;

Query Match 31.7%; Score 487; DB 24; Length 18596;

Best Local Similarity 99.7%; Pred. No. 4,66-241;

Matches 587; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 908 AGCTGAGGAGAACCCAGACCTTCCCAAGCTCAGATTCTTGGAAAAGTTAGAAAA 967
 Db 15611 AGCTTACGGAGAACCCAGACCTTCCCAAGCTCAGATTCTTGGAAAAGTTAGAAAA 15670
 QY 968 TTGATGACTTCAAACTGAGACCTTCAATTGAGGTTACATCCGATCAGATTATTA 1027
 Db 15671 TTGATGACTTCAAACTGAGACCTTCAATTGAGGTTACATCCGATCAGATTATTA 15730
 QY 1028 AATATGAAATGGCTGTTTAAAGGCTCTTCAAGAGCTTGAAGATATGTGCTCTTAA 1087
 Db 15731 AATATGAAATGGCTGTTTAAAGGCTCTTCAAGAGCTTGAAGATATGTGCTCTTAA 15790
 QY 1088 GGGGTGGGCTGGATGCCAGGTAAGTCTTTGCTTAAAGAAAGAAAGAACTAG 1147
 Db 15791 GGGGTGGGCTGGATGCCAGGTAAGTCTTTGCTTAAAGAAAGAAAGAACTAG 15850
 QY 1148 TCAAAAATCGTCCGAGACCTATGATTATTTTAAAGATGTTGCCACTGGCAAT 1207
 Db 15851 TCAAAAATCGTCCGAGACCTATGATTATTTTAAAGATGTTGCCACTGGCAAT 15910
 QY 1208 GTACTGTGCAAGTCTTTCATATATAAAGGCTTGAAGTAACTCACTGAGGATCTG 1267
 Db 15911 GTACTGTGCAAGTCTTTCATATATAAAGGCTTGAAGTAACTCACTGAGGATCTG 15970
 QY 1268 ACAATGCTGAGGTTATGAAACAAAGTGAAGAAATGATATGTGCTTTAGCAAAAAC 1327
 Db 15971 ACAATGCTGAGGTTATGAAACAAAGTGAAGAAATGATATGTGCTTTAGCAAAAAC 16030
 QY 1328 ATGATATGTCATTTCAATCCACGTACTTATTAAGAAGGTGTGTAATTTACAGACTAT 1387
 Db 16031 ATGATATGTCATTTCAATCCACGTACTTATTAAGAAGGTGTGTAATTTACAGACTAT 16090
 QY 1388 TTTTGAATTTTATTAAGATTTTATTAAGATTTTCAACAGCTTTCCCAATCTAGAGG 1447
 Db 16091 TTTTGAATTTTATTAAGATTTTATTAAGATTTTCAACAGCTTTCCCAATCTAGAGG 16150
 QY 1448 AGCTGAGTAAACCAATCATCATGATGATGATGATGATGATGATGATGATGATGATGAT 1496
 Db 16151 AGCTGAGTAAACCAATCATCATGATGATGATGATGATGATGATGATGATGATGATGAT 16199
 RESULT 10
 ABL63078

ID ABL63078 standard; DNA; 18596 BP.
XX
XX ABL63078;
XX
DT 15-MAY-2002 (first entry)
XX
DE Breast cancer related gene sequence SEQ ID NO:1415.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; carcinous;
KW cytostatic; gene therapy; antineoplastic; Wilms' tumour; adenocarcinoma;
XX gene; ds.
XX
OS Homo sapiens.
XX
PN WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US10838.
XX
PR 05-JUN-2000; 2000US-209473P.
PR 05-JUN-2000; 2000US-209531P.
PR 18-SEP-2000; 2000US-233133P.
PR 18-SEP-2000; 2000US-233617P.
PR 20-SEP-2000; 2000US-234009P.
PR 20-SEP-2000; 2000US-234034P.
PR 20-SEP-2000; 2000US-234052P.
PR 22-SEP-2000; 2000US-234509P.
PR 22-SEP-2000; 2000US-234567P.
PR 25-SEP-2000; 2000US-234923P.
PR 25-SEP-2000; 2000US-234924P.
PR 25-SEP-2000; 2000US-235077P.
PR 25-SEP-2000; 2000US-235082P.
PR 25-SEP-2000; 2000US-235134P.
PR 25-SEP-2000; 2000US-235280P.
PR 26-SEP-2000; 2000US-235637P.
PR 26-SEP-2000; 2000US-235638P.
PR 27-SEP-2000; 2000US-235711P.
PR 27-SEP-2000; 2000US-235720P.
PR 27-SEP-2000; 2000US-235840P.
PR 27-SEP-2000; 2000US-235863P.
PR 28-SEP-2000; 2000US-236028P.
PR 28-SEP-2000; 2000US-236032P.
PR 28-SEP-2000; 2000US-236033P.
PR 28-SEP-2000; 2000US-236034P.
PR 28-SEP-2000; 2000US-236109P.
PR 28-SEP-2000; 2000US-236111P.
PR 29-SEP-2000; 2000US-236842P.
PR 29-SEP-2000; 2000US-236881P.
PR 02-OCT-2000; 2000US-237172P.
PR 02-OCT-2000; 2000US-237173P.
PR 02-OCT-2000; 2000US-237278P.
PR 02-OCT-2000; 2000US-237284P.
PR 02-OCT-2000; 2000US-237285P.
PR 02-OCT-2000; 2000US-237316P.
PR 03-OCT-2000; 2000US-237455P.
PR 03-OCT-2000; 2000US-237582P.
PR 03-OCT-2000; 2000US-237604P.
PR 03-OCT-2000; 2000US-237606P.
PR 03-OCT-2000; 2000US-237608P.
PR 01-NOV-2000; 2000US-244867P.
PR 01-NOV-2000; 2000US-245084P.
XX
XX (AVAL-) AVALON PHARM.
XX
XX Young PG, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S,
PI Sopot DR, Weaver Z;
XX
XX WPI; 2002-188264/24.
XX
PT Screening for anti-neoplastic agent involves exposing cells to a
PT chemical agent to be tested for anti-neoplastic activity; and

PT determining a change in expression of a gene of a signature gene set -
XX
XX Claim 1; SEQ ID 1415; 44bp; English.
XX
CC The present invention describes a method (M1) for screening for an
CC anti-neoplastic agent. The method involves exposing cells to a chemical
CC agent to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL0110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening
CC an anti-neoplastic agent, and can be used for producing a product which
CC is the data collected with respect to the anti-neoplastic agent as a
CC result of M1, and the data is sufficient to convey the chemical
CC structure and/or properties of the agent. M1 can be used in the
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer.
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC carcinoma, papillary carcinoma and Wilms' tumour.
XX
SQ Sequence 18596 BP; 4521 A; 3991 C; 4479 G; 5605 T; 0 other;
Query Match 31.7%; Score 487; DB 24; Length 18596;
Best Local Similarity 99.7%; Pred. No. 4.6e-241;
Matches 587; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 908 AGCTGAGCGAGAACCCAGACCTTCCCAAGCTCAGAGATTCTGGAAGAAGTGAAGAA 967
DB 15611 AGCTGAGCGAGAACCCAGACCTTCCCAAGCTCAGAGATTCTGGAAGAAGTGAAGAA 15670
QY 968 TTGATGACTTCAAGCTGAGAACTTTCAGATTGAGAGGTTCATCCGATCCATCTATTA 1027
DB 15671 TTGATGACTTCAAGCTGAGAACTTTCAGATTGAGAGGTTCATCCGATCCATCTATTA 15730
QY 1028 AATGGAATGCTGTTTAAAGGTCTTCAAGAGCTGAGAGATATGTCAGTCTTTA 1087
DB 15731 AATGGAATGCTGTTTAAAGGTCTTCAAGAGCTGAGAGATATGTCAGTCTTTA 15790
QY 1088 GGGGTGGGCTGAGTCCAGGTAAAGTCTTTTCTCTTAAAGAAAGAAAGAACTGAG 1147
DB 15791 GGGGTGGGCTGAGTCCAGGTAAAGTCTTTTCTCTTAAAGAAAGAAAGAAAGTGG 15850
QY 1148 TCAAAATCTGTCGCGACCTATGCTATTAATTTTAAAGATGTCGCACTGGCAAT 1207
DB 15851 TCAAAATCTGTCGCGACCTATGCTATTAATTTTAAAGATGTCGCACTGGCAAT 15910
QY 1208 GTAATGCGAGTCTTCCATTAATAAAGCTTTGATTACTCAGTGAAGGATCTG 1267
DB 15911 GTAATGCGAGTCTTCCATTAATAAAGCTTTGATTACTCAGTGAAGGATCTG 15970
QY 1268 ACAATGCTGAGTTATGAACAAGTGAAGAAATGATGTCCTTGAAGAAAAAC 1327
DB 15971 ACAATGCTGAGTTATGAACAAGTGAAGAAATGATGTCCTTGAAGAAAAAC 16030
QY 1328 ATGATGTCATTTCAATCCACGTAATTAATAAAGTGGTGATTTTCAAGCTAT 1387
DB 16031 ATGATGTCATTTCAATCCACGTAATTAATAAAGTGGTGATTTTCAAGCTAT 16090
QY 1388 TTTTGGAAATTTTGAATATTTTAAAGATTTTCAAGCTATTTCCCAATCTGAGGG 1447
DB 16091 TTTTGGAAATTTTGAATATTTTAAAGATTTTCAAGCTATTTCCCAATCTGAGGG 16150
QY 1448 AGCTGAGTAACACCATGATCATGATGATGAGAGTGGTATGATGACTTAA 1496
DB 16151 AGCTGAGTAACACCATGATCATGATGATGAGAGTGGTATGATGACTTAA 16199
RESULT 11
ID ABL67927
XX
XX ABL67927 standard; DNA; 18596 BP.
XX
PT

AC ABL67927;
 XX
 DT 15-MAY-2002 (first entry)
 XX
 DE Ovary cancer related gene sequence SEQ ID NO:6264.
 XX
 KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
 KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
 KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
 KW gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN M0200194629-A2.
 XX
 PD 13-DEC-2001.
 XX
 PF 30-MAY-2001, 2001WO-US10838.
 XX
 PR 05-JUN-2000, 2000US-209473P.
 PR 05-JUN-2000, 2000US-209531P.
 PR 18-SEP-2000, 2000US-231133P.
 PR 18-SEP-2000, 2000US-233617P.
 PR 20-SEP-2000, 2000US-234009P.
 PR 20-SEP-2000, 2000US-234034P.
 PR 20-SEP-2000, 2000US-234052P.
 PR 22-SEP-2000, 2000US-234509P.
 PR 22-SEP-2000, 2000US-234567P.
 PR 25-SEP-2000, 2000US-234923P.
 PR 25-SEP-2000, 2000US-234924P.
 PR 25-SEP-2000, 2000US-235077P.
 PR 25-SEP-2000, 2000US-235082P.
 PR 25-SEP-2000, 2000US-235134P.
 PR 25-SEP-2000, 2000US-235280P.
 PR 26-SEP-2000, 2000US-235637P.
 PR 26-SEP-2000, 2000US-235638P.
 PR 27-SEP-2000, 2000US-235711P.
 PR 27-SEP-2000, 2000US-235720P.
 PR 27-SEP-2000, 2000US-235840P.
 PR 27-SEP-2000, 2000US-235863P.
 PR 28-SEP-2000, 2000US-236028P.
 PR 28-SEP-2000, 2000US-236032P.
 PR 28-SEP-2000, 2000US-236033P.
 PR 28-SEP-2000, 2000US-236034P.
 PR 28-SEP-2000, 2000US-236109P.
 PR 28-SEP-2000, 2000US-236111P.
 PR 29-SEP-2000, 2000US-236842P.
 PR 29-SEP-2000, 2000US-236891P.
 PR 02-OCT-2000, 2000US-237172P.
 PR 02-OCT-2000, 2000US-237173P.
 PR 02-OCT-2000, 2000US-237278P.
 PR 02-OCT-2000, 2000US-237295P.
 PR 02-OCT-2000, 2000US-237316P.
 PR 03-OCT-2000, 2000US-237425P.
 PR 03-OCT-2000, 2000US-237598P.
 PR 03-OCT-2000, 2000US-237604P.
 PR 03-OCT-2000, 2000US-237606P.
 PR 03-OCT-2000, 2000US-237608P.
 PR 01-NOV-2000, 2000US-244867P.
 PR 01-NOV-2000, 2000US-245084P.
 XX
 PA (AVAL-) AVALON PHARM.
 XX
 PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
 PI Soppet DR, Weaver Z;
 XX
 DR WPI; 2002-188264/24.
 XX
 PT Screening for anti-neoplastic agent involves exposing cells to a
 PT chemical agent to be tested for anti-neoplastic activity, and
 PT determining a change in expression of a gene of a signature gene set -
 XX

PS Claim 1: SEQ ID 6264; 449P; English.
 XX
 CC The present invention describes a method (M1) for screening for an
 CC anti-neoplastic agent. The method involves exposing cells to a chemical
 CC agent to be tested for anti-neoplastic activity, determining a change in
 CC expression of at least one gene (I) of a signature gene set, where (I)
 CC comprises a sequence (S) selected from 847 sequences (given in ABL61664
 CC to ABL70110), or is at least 95% identical to (S), where a change in
 CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
 CC activity and can be used in gene therapy. M1 can be used for screening
 CC an anti-neoplastic agent, and can be used for producing a product which
 CC is the data collected with respect to the anti-neoplastic agent as a
 CC result of M1, and the data is sufficient to convey the chemical
 CC structure and/or properties of the agent. M1 can be used in the
 CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
 CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
 CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
 CC carcinoma, papillary carcinoma and Wilm's tumour.
 XX
 SQ Sequence 18596 BP; 4521 A; 3991 C; 4479 G; 5605 T; 0 other;
 XX
 Query Match 31.7%; Score 487; DB 24; Length 18596;
 Best Local Similarity 99.7%; Pred. No. 4,5e-241;
 Matches 587; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 XX
 QY 908 AGCTTACGAGGAAACCCAGACCTTCCCAAGCTCAGGATCTTCGAAAGTGAAGAAA 967
 DB 15611 AGCTTACGAGGAAACCCAGACCTTCCCAAGCTCAGGATCTTCGAAAGTGAAGAAA 15670
 QY 968 TTGATGACTTCAAGCTGAGAGAGCTTCAAGTTGAAGGAGTACATCCGATCACTATTA 1027
 DB 15671 TTGATGACTTCAAGCTGAGAGAGCTTCAAGTTGAAGGAGTACATCCGATCACTATTA 15730
 QY 1028 AATGGAATGAGCTGTTAGGGTCTTTCAGAGAGCTGAGAGATGTCAGTCTTTA 1087
 DB 15731 AATGGAATGAGCTGTTAGGGTCTTTCAGAGAGCTGAGAGATGTCAGTCTTTA 15790
 QY 1088 GGGGTGGGCTGAGATGCCAGGTAAGAAAGTCTTTTGTCTTAAGAAAGAAAGAACTAGG 1147
 DB 15791 GGGGTGGGCTGAGATGCCAGGTAAGAAAGTCTTTTGTCTTAAGAAAGAAAGAACTAGG 15850
 QY 1148 TCAAAATCTGCGCTGACCTATGATGTTATATTTTAAAGATGTTCCACGCGAAT 1207
 DB 15851 TCAAAATCTGCGCTGACCTATGATGTTATATTTTAAAGATGTTCCACGCGAAT 15910
 QY 1208 GTACCTGTCGAGTCTTTCATATATTAAGGCTTGAAGTACCTGAGGGTATCTG 1267
 DB 15911 GTACCTGTCGAGTCTTTCATATATTAAGGCTTGAAGTACCTGAGGGTATCTG 15970
 QY 1268 ACAATGCTGAGGTTATGAACAAAGTGAAGAAATGATGCTTACGAAATAAC 1327
 DB 15971 ACAATGCTGAGGTTATGAACAAAGTGAAGAAATGATGCTTACGAAATAAC 16030
 QY 1328 ATGATGCTGATTTCAATCCAGTACTTAAAGAGGTTGTGAATTTTCAACACTAT 1387
 DB 16031 ATGATGCTGATTTCAATCCAGTACTTAAAGAGGTTGTGAATTTTCAACACTAT 16090
 QY 1388 TTTTGAATATTTTGAATATTTTAAAGAAATTTCAACAGCTTATCCCTCAATCTGAGG 1447
 DB 16091 TTTTGAATATTTTGAATATTTTAAAGAAATTTCAACAGCTTATCCCTCAATCTGAGG 16150
 QY 1448 AGCTGAGTACACCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1496
 DB 16151 AGCTGAGTACACCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 16199
 RESULT 12
 ID ABA93401/c
 ID ABA93401 standard; DNA; 45716 BP.
 AC ABA93401;
 XX

Oy	1148	TCAAAATCTGTCGGTGAACCTACATGTTATTAATTTTAAAGATGTGCCACTGGCAAT	1207
Db	44342	TCAAAATCTGTCGGTGAACCTACATGTTATTAATTTTAAAGATGTGCCACTGGCAAT	4428
Oy	1208	GTACCTGTGCCAGTTCCTTCATATAAAGGCTTTGAGTTACTCACTGAGGATCTTG	1267
Db	44282	GTACCTGTGCCAGTTCCTTCATATAAAGGCTTTGAGTTACTCACTGAGGATCTTG	4422
Oy	1268	ACAATCTGAGGTTATGAAACAAAGTAGAGAAATGAATGTATGTGCTCTTTCGAAAC	1327
Db	44222	ACAATCTGAGGTTATGAAACAAAGTAGAGAAATGAATGTATGTGCTCTTTCGAAAC	4416
Oy	1328	ATATATGTGCAATTCACATCCCAAGTACTATAAAGAGTTGGTAATTCACAAGCTAT	1387
Db	44162	ATATATGTGCAATTCACATCCCAAGTACTATAAAGAGTTGGTAATTCACAAGCTAT	4410
Oy	1388	TTTTGGAAATATTTTAAATATTTTAAAGATTTTCAAAAGCTATTCCTCAAACTGAGG	1447
Db	44102	TTTTGGAAATATTTTAAAGATATTTTAAAGATTTTCAAAAGCTATTCCTCAAACTGAGG	4404
Oy	1448	AGCTGATGTAACCACTGATCATGATGTATGAGGTGGTTATGAACCTTA	1496
Db	44042	AGCTGATGTAACCACTGATCATGATGTATGAGGTGGTTATGAACCTTA	43994
RESULT 13			
ABA93402/c	ABA93402/c		
ID	ABA93402 standard; DNA; 45989 BP.		
AC	ABA93402;		
DT	22-APR-2002 (first entry)		
DE	Human rTS-beta gene SEQ ID NO:6.		
KM	Human: rTS-alpha; rTS-beta; rTS; thymidylate synthase; chromosome 18;		
KM	mutant thymidylate synthase; enzyme; antineoplastic; antidepressant;		
KM	neuroleptic; nootropic; tranquilizer; gene therapy; hypomania;		
KM	neuropsychiatric disorder; bipolar affective disorder; schizophrenia;		
KM	severe bipolar affective disorder; bipolar affective disorder;		
KM	major depression; attention deficit disorder; schizoaffective disorder		
OS	Homo sapiens.		
PN	US6323244-B1.		
FD	27-NOV-2001.		
XX	05-NOV-1997; 97US-0965048.		
PR	05-NOV-1997; 97US-0965048.		
PA	(MILL-) MILLENNIUM PHARM INC.		
PA	(REGC) UNIV CALIFORNIA.		
PI	Chen H, Freimer NB;		
DR	WEI; 2002-112960/15.		
PT	Treating an rTS (mutant form of thymidylate synthase) - mediated		
PT	neuropsychiatric disorder, e.g., bipolar affective disorder, comprises		
PT	administering a compound, identified by an rTS assay, that interferes		
PT	with rTS gene product -		
XX	Disclosure; Fig 4A-S; 110p; English.		
CC	The present invention describes a method for treating an rTS (a mutant		
CC	form of thymidylate synthase) mediated neuropsychiatric disorder. The		
CC	method comprises identifying a compound (C) that can be used to treat		
CC	the disorder, by contracting a small molecule test compound (I) with an		
CC	rTS protein (II) comprising a sequence of either 361 or 416 amino acids		

CC (AB05587 or AB05588) determining whether (I) binds to (II). (C) has
 CC antianemic, antidepressant, neuroleptic, nootropic and tranquilizer
 CC activities. rts polynucleotide sequences can be used in gene therapy.
 CC The method is useful in identifying compounds for treating rts-mediated
 CC neuropsychiatric disorders including bipolar affective disorder,
 CC e.g., severe bipolar affective (mood) disorder, bipolar affective (mood)
 CC disorder with hypomania, major depression, schizophrenia, attention
 CC deficit disorder and schizoaffective disorder. The present sequence
 CC represents human rts-beta gene which is used in the exemplification of
 CC the present invention. rts has been located to chromosome 18.

XX Sequence 45989 BP; 12183 A; 10457 C; 10488 G; 12799 T; 62 other;

Query Match 31.7%; Score 487; DB 24; Length 45989;
 Best Local Similarity 99.7%; Pred. No. 4.8e-241;
 Matches 587; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 908 AGCTTCACGGAAGACCCGACCTTTCCCAAGCTCAGATTCTTGGAAAAGTTGAGAAA 967
 DB 44855 AGCTTCACGGAAGACCCGACCTTTCCCAAGCTCAGATTCTTGGAAAAGTTGAGAAA 44796
 QY 968 TTGATGACTTCAAGAGCTGAAGAGCTTCAGATTGAAGAGGTACATCCGATCCACTATTA 1027
 DB 44795 TTGATGACTTCAAGAGCTGAAGAGCTTCAGATTGAAGAGGTACATCCGATCCACTATTA 44736
 QY 1028 AATGGAATGCTGTTTAAAGGCTTCGAAAGGCTGAGAGATTTGTGAGCTTTA 1087
 DB 44735 AATGGAATGCTGTTTAAAGGCTTCGAAAGGCTGAGAGATTTGTGAGCTTTA 44676
 QY 1088 GGGGTGGGCTGAGTGCAGAGTGAAGATTTCTTTTGTCTAAAGAGAAAGAACTAGG 1147
 DB 44675 GGGGTGGGCTGAGTGCAGAGTGAAGATTTCTTTTGTCTAAAGAGAAAGAACTAGG 44616
 QY 1148 TCAAAATCTGTCCGATGACCTATCACTATTAATTTTAAAGATGTTGCCACTGCAAT 1207
 DB 44615 TCAAAATCTGTCCGATGACCTATCACTATTAATTTTAAAGATGTTGCCACTGCAAT 44556
 QY 1208 GTACTGTGCGAGTTCTTTCATTAATAAAGGCTTTGATTACTCTGAGGGTATCTG 1267
 DB 44555 GTACTGTGCGAGTTCTTTCATTAATAAAGGCTTTGATTACTCTGAGGGTATCTG 44496
 QY 1268 ACAATGCTGAGTTATGAAGAGTGAAGAGATTAATGATGATGCTCTGAGCAAAAC 1327
 DB 44495 ACAATGCTGAGTTATGAAGAGTGAAGAGATTAATGATGATGCTCTGAGCAAAAC 44436
 QY 1328 ATGTATGTCATTTCAATCCAGTACTTAATAAGAGTGTGGAATTTCAAGACTAT 1387
 DB 44435 ATGTATGTCATTTCAATCCAGTACTTAATAAGAGTGTGGAATTTCAAGACTAT 44376
 QY 1388 TTTTGAATATTTTAAAGATTTTAAAGATTTTCAAGATTTTCCCTCAATCTGAGG 1447
 DB 44375 TTTTGAATATTTTAAAGATTTTAAAGATTTTCAAGATTTTCCCTCAATCTGAGG 44316
 QY 1448 AGCTGATGAACACCATGATCATGATGATGATGATGATGATGATGATGATGATGAT 1496
 DB 44315 AGCTGATGAACACCATGATCATGATGATGATGATGATGATGATGATGATGATGAT 44267

RESULT 14
 AAX24270
 ID AAX24270 standard; DNA; 566 BP.
 XX AAX24270;
 AC
 XX 02-JUL-1999 (first entry)
 DT
 XX Human thymidylate synthase DNA.
 DE
 XX Thymidylate synthase; antisense oligonucleotide; cytosstatic;
 KW tumor cell; anticancer drug; treatment; cancer; antiproliferative;
 KW antioestrogen; progestogen; antiandrogen; testosterone inhibitor;
 KW anti-invasion agent; growth factor inhibitor; antimetabolite;
 KW antibiotic; alkylating agent; antimetabolic agent; radiotherapy;

KW topoisomerase inhibitor; ss.
 XX Homo sapiens.
 OS
 XX MO9915648-A1.
 PN
 XX 01-APR-1999.
 PD
 XX 17-SEP-1998; 98MO-GB02820.
 PF
 XX 06-JUN-1998; 98GB-0012140.
 PR 23-SEP-1997; 97GB-0020107.
 PR 17-OCT-1997; 97GB-0022012.
 XX
 PA (ISIS-) ISIS PHARM INC.
 PA (ZENB-) ZENBECA LTD.
 XX
 PI Dean NM, Koropatnick DJ, Vincent MD;
 XX
 XX WPI; 1999-254708/21.
 DR
 XX
 XX
 PS
 PS Disclosure; Fig 7, 53pp; English.
 CC This invention describes novel antisense oligonucleotides targeted to
 CC sequences in the 3' end of thymidylate synthase (TS) mRNA. Such
 CC oligonucleotides are cytostatic on their own when administered to
 CC human tumor cell lines, and also enhance the toxicity of anticancer
 CC drugs such as Tomudex administered to those cells. In addition, antisense
 CC oligonucleotides targeted to 5' sequences induce TS gene transcription.
 CC The antisense oligonucleotides are used in a method for the
 CC treatment of cancer (or a method for providing antiproliferative effect)
 CC The antisense oligonucleotides are also used in the production of a
 CC medicament for the treatment of cancer, either separately or in conjunction
 CC with a therapeutic agent such as thymidylate synthase (TS) inhibitors
 CC (e.g. Tomudex, Zeneca development compound ZD9331 etc.); cytosstatic
 CC agents (e.g. antioestrogens, (anti)progestogens, antiandrogens,
 CC testosterone inhibitors, anti-invasion agents, growth factor inhibitors,
 CC etc.); antiproliferative/antineoplastic agents (e.g. antimetabolites,
 CC antitumor antibiotics, alkylating agents, antimetabolic agents,
 CC topoisomerase inhibitors, etc.); or radiotherapy.
 CC
 XX
 SQ Sequence 566 BP; 156 A; 101 C; 135 G; 174 T; 0 other;

Query Match 24.7%; Score 379; DB 20; Length 566;
 Best Local Similarity 99.8%; Pred. No. 2.8e-185;
 Matches 429; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1067 GAAGATATTTGCACTTTAGGGGTTGGGCTGATGCCGAGTAAAGTCTTTTGTCT 1126
 DB 97 GAAGATATTTGCACTTTAGGGGTTGGGCTGATGCCGAGTAAAGTCTTTTGTCT 156
 QY 1127 CTAAAGAGAAAGAGACTAGGTCAAAAATCTGTCCGTGACCTTCACTATTAATTTTA 1186
 DB 157 CTAAAGAGAAAGAGACTAGGTCAAAAATCTGTCCGTGACCTTCACTATTAATTTTA 216
 QY 1187 AGGATGTGCGACCTGGCAAAATGTAACGTGACAGTCTTTCATTAATAAAGGCTTGGAG 1246
 DB 217 AGGATGTGCGACCTGGCAAAATGTAACGTGACAGTCTTTCATTAATAAAGGCTTGGAG 276
 QY 1247 TTAACCTGAGGATATGACAAATGCTGAGATTATGAACAAAGTGAAGAGATGAAT 1306
 DB 277 TTAACCTGAGGATATGACAAATGCTGAGATTATGAACAAAGTGAAGAGATGAAT 336
 QY 1307 GTATGTGCTTTAGCAAAAACATGATATGCAATTTCAATCCAGTACTTATAAAGAG 1366
 DB 337 GTATGTGCTTTAGCAAAAACATGATATGCAATTTCAATCCAGTACTTATAAAGAG 396
 QY 1367 TTGATGAATTTACAAAGTATTTTGGAAATTTTAAAGATTTTAAAGATTTTCAAG 1426
 DB 397 TTGATGAATTTACAAAGTATTTTGGAAATTTTAAAGATTTTAAAGATTTTCAAG 456

QY 1427 CTATTCCTCAATCTGAGGAGCTGAGTAACACCATCGATCGATGATGATGAGTGGTT 1486
 DB 457 CTATTCCTCAATCTGAGGAGCTGAGTAACACCATCGATCGATGATGATGAGTGGTT 516
 QY 1487 ATGAACCTTTA 1496
 DB 517 ATGAACCTTTA 526

RESULT 15

AA884956
 ID AA884956 standard; cDNA; 651 BP.

AC AA884956;

DT 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #20760.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 food supplement; medical imaging; diagnostic; genetic disorder; ss.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI: 2001-639362/73.

DR P-PSDB; ABG20769.

PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity

PS Claim 1; SEQ ID No 20760; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AA864197-AA894564 represent novel human
 CC diagnostic coding sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

SO Sequence 651 BP; 152 A; 86 C; 241 G; 172 T; 0 other;

Query Match 11.6%; Score 178; DB 23; Length 651;
 Best Local Similarity 100.0%; Pred. No. 2.2e-81;

Matches 178; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 661 GATCTTCTCTGATGAGGCTGACCTCCATGACCATGACCTCTGACGATTCATGAGTGAAC 720
 DB 337 GATCTTCTCTGATGAGGCTGACCTCCATGACCATGACCTCTGACGATTCATGAGTGAAC 396
 QY 721 ACTGAGCTGTCTGACGCTGATGACGAGATCGGAGACATGGGCTCTGAGCTTTTC 780
 DB 397 ACTGAGCTGTCTGACGCTGATGACGAGATCGGAGACATGGGCTCTGAGCTTTTC 456
 QY 781 AACATCGCAGCTACGCGCTGCTCAAGTACATGATTTGGACATCAAGGCGCTGAAGC 838
 DB 457 AACATCGCAGCTACGCGCTGCTCAAGTACATGATTTGGACATCAAGGCGCTGAAGC 514

Search completed: November 29, 2002, 00:34:55
 Job time : 468.605 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 28, 2002, 19:30:54 ; Search time 1563.2 Seconds
(without alignments)
12297.911 Million cell updates/sec

Title: US-09-963-333-7

Perfect score: 1187

Sequence: 1 gatcgccacccacccaccca.....gagggagcgccgctcgggga 1187

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estbda:*
2: em_esthum:*
3: em_estlin:*
4: em_estm:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_estl:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estum:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match Length	DB	ID	Description
1	367.6	31.0	1061	14	BQ943386 AGENCOURT
2	337.4	28.4	1060	13	BM456638 AGENCOURT
3	331.4	27.9	741	9	AUI18644 AUI18644
4	329.8	27.8	739	9	AUI43295 AUI43295
5	329.8	27.8	819	9	AUI43180 AUI43180
6	329.4	27.8	786	9	AL545940 AL545940

7	325.6	27.4	1197	14	BQ072305
8	322.2	27.1	1622	14	BM917794 AGENCOURT
9	320.4	27.0	841	14	BQ34233 AGENCOURT
10	317.4	26.7	985	13	BM466216
11	317.2	26.7	1018	14	BM600922
12	312.2	26.3	1021	14	BQ685120
13	312.2	26.3	668	9	AL541351
14	305.4	25.7	938	9	AL551990
15	298.4	25.1	523	14	BM763658
16	297.4	25.1	738	9	AUI30205
17	295.8	24.9	711	9	AUI42980
18	294.2	24.8	531	14	BM746027
19	290.2	24.4	526	9	AUI28726
20	288.4	24.3	529	12	BM794349
21	287.6	24.2	905	13	BIS97722
22	287.4	23.8	948	14	BQ066205
23	281.8	23.7	935	9	AL542409
24	281.2	23.7	524	12	BQ482405
25	280.4	23.6	1045	13	BM478045
26	278.6	23.5	678	12	BG337663
27	278.6	23.5	832	12	BF685317
28	278.4	23.5	742	9	AL548721
29	278.4	23.5	885	14	BQ424746
30	277.8	23.4	1055	14	BQ86916
31	277	23.3	970	14	BQ928426
32	276.8	23.3	1020	9	AL547612
33	275.8	23.2	883	14	BQ420981
34	268.4	22.6	994	14	BM304778
35	268.2	22.6	664	13	BIS20187
36	267.8	22.6	406	14	BM834317
37	267.8	22.6	1149	13	BM546045
38	267.6	22.5	963	12	BG337540
39	267.4	22.5	809	9	AUI117646
40	267	22.5	267	13	BI668571
41	266.8	22.5	538	13	BM41631
42	266.6	22.5	439	12	BF689564
43	265.4	22.4	672	12	BG574267
44	264.6	22.3	950	9	AL550812
45	264.4	22.3	688	9	AL543217

ALIGNMENTS

RESULT 1
BQ943386 1061 bp mRNA linear EST 21-AUG-2002
LOCUS AGENCOURT 8804318 NIH_MGC_47 Homo sapiens cDNA clone IMAGE:6377891
DEFINITION 5', mRNA sequence.
ACCESSION BQ943386
VERSION BQ943386
KEYWORDS EST.
SOURCE BQ943386.1 GI:22358676
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 1061)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at:
<http://image.llnl.gov>
Plate: LNCM562 row: g column: 12
High quality sequence steps: 382.

FEATURES

source 1. 1061

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:6377891"
 /clone_1ib="NIH_MGC_47"
 /tissue_type="neuroblastoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: brain; Vector: pOTB7; Site: 1: XhoI; Site: 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library."
 BASE COUNT 202 a 402 c 285 g 171 t 1 others
 ORIGIN

Query Match 31.0%; Score 367.6; DB 14; Length 1061;
 Best Local Similarity 92.7%; Pred. No. 3e-36; 4; Indels 28; Gaps 1;
 Matches 408; Conservative 0; Mismatches 4; Indels 28; Gaps 1;

QY 694 GTGCCACACCCGCTGCTCTGCTTCCCTCGGCGCACGCTCTTGAAGCGGAGCGCGC
 Db 1 GTGCCACACCCGCTGCTCTGCTTCCCTCGGCGCACGCTCTTGAAGCGGAGCGCGC 60
 QY 754 CGCGACCCCGCGCGAGAGAGAGCGAGCGCGCGCGGAGAGAGCGCGCG 813
 Db 61 CGCGACCCCGCGCGAGAGAGAGCGAGCGCGCGCGGAGAGAGCGCGCG 120
 QY 814 AAGGGGTCTTGGCACCAGGCTTGGCTCTGCTTCCGCGCGCGCGCGCGCGCG 873
 Db 121 AAGGGGTCTTGGCACCAGGCTTGGCTCTGCTTCCGCGCGCGCGCGCGCGCG 152
 QY 874 CTCGCTCCCG 933
 Db 153 CTCGCTCCCG 212
 QY 934 GCGGCGCTGAGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 993
 Db 213 GCGGCGCTGAGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 272
 QY 994 CCGCGCTCCCGCGCGAGCGAGCTGAGTACTTGGGAGCATCAACATCTCTCGCTGC 1053
 Db 273 CCGCGCTCCCGCGAGCGAGCTGAGTACTTGGGAGCATCAACATCTCTCGCTGC 332
 QY 1054 GCGCGTCAAGAGAGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCG 1113
 Db 333 GCGCGTCAAGAGAGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCG 392
 QY 1114 CGCTACAGCTGAGAGGTGA 1133
 Db 393 CGCTACAGCTGAGAGGTGA 412

RESULT 2
 BM456638 1060 bp mRNA linear EST 05-FEB-2002
 LOCUS BM456638
 DEFINITION AGSCOURT 6408748 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:5496315
 5', mRNA sequence.
 ACCESSION BM456638
 VERSION BM456638.1 GI:18505678
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1060)
 AUTHORS NIH-MGC http://mgs.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-r@mail.nih.gov

Tissue Procurement: Lou Straudt
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LIML at:
 http://image.llnl.gov
 Plate: LIML2124 row: k column: 04
 High quality sequence stop: 530.
 Location/Qualifiers
 1. 1060
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5496315"
 /clone_1ib="NIH_MGC_85"
 /tissue_type="lymphoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: lymph; Vector: pCMV-SPORT6; Site: 1: NotI; Site: 2: SalI; Cloned unidirectionally, oligo-dT primed. Average insert size 1.867 kb. Library enriched for full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."

BASE COUNT 217 a 396 c 259 g 188 t
 ORIGIN

Query Match 28.4%; Score 337.4; DB 13; Length 1060;
 Best Local Similarity 99.7%; Pred. No. 1.4e-32;
 Matches 338; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 795 CGCGGAGAGAGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 854
 Db 1 CGCGGAGAGAGAGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60
 QY 855 GCGCGCGCATCTTGGCTGCTCTGCTCCGCGCGCGCGCGCATCTTGGCTGCTCCGCG 914
 Db 61 GCGCGCGCATCTTGGCTGCTCTGCTCCGCGCGCGCGCGCATCTTGGCTGCTCCGCG 120
 QY 915 CGCGCGCGCATCTTGGCTGCTCTGCTCCGCGCGCGCGCGCATCTTGGCTGCTCCGCG 974
 Db 121 CGCGCGCGCATCTTGGCTGCTCTGCTCCGCGCGCGCGCGCATCTTGGCTGCTCCGCG 180
 QY 975 ACGAGAGCGAGAGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1034
 Db 181 ACGAGAGCGAGAGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 240
 QY 1035 CCAACATCTCTCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1094
 Db 241 CCAACATCTCTCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
 QY 1095 GGTATTGCGCATGAGAGCGCGCTTACAGCTTGAAGGTGA 1133
 Db 301 GGTATTGCGCATGAGAGCGCGCTTACAGCTTGAAGGTGA 339

RESULT 3
 AUI18644 741 bp mRNA linear EST 01-AUG-2002
 LOCUS AUI18644
 DEFINITION AUI18644 HEMBA1 Homo sapiens cDNA clone HEMBA1004057 5', mRNA
 sequence.
 ACCESSION AUI18644
 VERSION AUI18644.1 GI:10933771
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 741)
 AUTHORS Oca, T., Nishikawa, T., Suzuki, Y., Ishii, S., Saito, K., Kawai, Y., Yamamoto, Y., Wakamatsu, A., Nakamura, Y., Nagai, T., Sugano, S. and Isogai, T.
 TITLE HRI human cDNA project
 JOURNAL Unpublished (2000)
 COMMENT Contact: Takao Isogai

Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp

HRI human CDNA project; 5'- & 3'-end one pass sequencing: Helix Research Institute; cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.

FEATURES

source

Location/Qualifiers

1..741

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="HEMBA1004057"

/clone_id="HEMBA1"

/tissue type="whole embryo, mainly head"

/dev stage="embryo, 10 weeks"

/note="Vector: pME18SFL3"

BASE COUNT 153 a 227 c 218 g 140 t 3 others

ORIGIN

Query Match

27.9%; Score 331.4; DB 9; Length 741;

Best Local Similarity 99.7%; Pred. NO. 9.9e-32;

Matches 332; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 801 AAAAGGCGCGGAGAGGGTCTGCGACCGCGCCACTTGAGCTGCTCCGCGCGCG 860
DB 2 AAAAGGCGCGGAGAGGGTCTGCGACCGCGCCACTTGAGCTGCTCCGCGCGCG 61
QY 861 CCACTTGAGCTGCTCCGCGCGCGCGCCACTTGAGCTGCTCCGCGCGCGCG 920
DB 62 CCACTTGAGCTGCTCCGCGCGCGCGCCACTTGAGCTGCTCCGCGCGCGCG 121
QY 921 CGCCATGCTGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 980
DB 122 CGCCATGCTGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 181
QY 981 GCGGAGCGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1040
DB 182 GCGGAGCGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 241
QY 1041 CATCTTCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1100
DB 242 CATCTTCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 301
QY 1101 CGGATGCGAGCGCGCTACAGCTGAGAGTGA 1133
DB 302 CGGATGCGAGCGCGCTACAGCTGAGAGTGA 334

RESULT 4

AUI43295

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AUI43295 739 bp mRNA linear EST 05-AUG-2002
AUI43295 Y79AA1 Homo sapiens cDNA clone Y79AA1001699 5', mRNA
sequence.
AUI43295
AUI43295.1 GI:11004816
EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 739)
Ota, T., Nishikawa, T., Suzuki, Y., Ishii, S., Saito, K., Kawai, Y.,
Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagai, T., Sugano, S. and
Isogai, T.
HRI human CDNA project
Unpublished (2000)
Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan

FEATURES

source

Location/Qualifiers

1..739

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="Y79AA1001699"

/clone_id="Y79AA1"

/cell type="retinoblastoma"

/cell_line="Y79"

/note="Vector: pME18SFL3"

BASE COUNT 155 a 223 c 217 g 140 t 4 others

ORIGIN

Query Match

27.8%; Score 329.8; DB 9; Length 739;

Best Local Similarity 99.4%; Pred. NO. 1.6e-31;

Matches 331; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 801 AAAAGGCGCGGAGAGGGTCTGCGACCGCGCCACTTGAGCTGCTCCGCGCGCG 860
DB 2 AAAAGGCGCGGAGAGGGTCTGCGACCGCGCCACTTGAGCTGCTCCGCGCGCG 61
QY 861 CCACTTGAGCTGCTCCGCGCGCGCGCCACTTGAGCTGCTCCGCGCGCGCG 920
DB 62 CCACTTGAGCTGCTCCGCGCGCGCGCCACTTGAGCTGCTCCGCGCGCGCG 121
QY 921 CGCCATGCTGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 980
DB 122 CGCCATGCTGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 181
QY 981 GCGGAGCGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1040
DB 182 GCGGAGCGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 241
QY 1041 CATCTTCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1100
DB 242 CATCTTCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 301
QY 1101 CGGATGCGAGCGCGCTACAGCTGAGAGTGA 1133
DB 302 CGGATGCGAGCGCGCTACAGCTGAGAGTGA 334

RESULT 5

AUI43180

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AUI43180 819 bp mRNA linear EST 05-AUG-2002
AUI43180 Y79AA1 Homo sapiens cDNA clone Y79AA1001513 5', mRNA
sequence.
AUI43180
AUI43180.1 GI:11004701
EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 819)
Ota, T., Nishikawa, T., Suzuki, Y., Ishii, S., Saito, K., Kawai, Y.,
Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagai, T., Sugano, S. and
Isogai, T.
HRI human CDNA project
Unpublished (2000)
Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp

HRI human cDNA project; 5'- & 3'-end one pass sequencing; Helix Research Institute; cDNA library construction; Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.

FEATURES

source

Location/Qualifiers
 1. .819
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="HY9AA1001513"
 /clone_lib="HY9AA1"
 /cell_type="retinoblastoma"
 /cell_line="Y79"
 /note="Vector: pME18SFL3"
 BASE COUNT 169 a 243 c 239 g 164 t 4 others
 ORIGIN

Query Match 27.8%; Score 329.8; DB 9; Length 819;
 Best Local Similarity 99.4%; Pred. No. 1.4e-31;
 Matches 331; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

801 AAAAGGCGCGGAGAGGGGTCTGCGACCGGCGCACTTGGCTGCTCCGCGCG 860
 2 AAAAGGCGCGGAGAGGGGTCTGCGACCGGCGCACTTGGCTGCTCCGCGCG 61
 861 CCACTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 920
 62 CCACTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 121
 921 CGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 980
 122 CGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 101
 981 GCGGAGCGCGAGCGCGCTGCGCGCGCGAGCTGCACTTGGGCGAGATCAACA 1040
 182 GCGGAGCGCGAGCGCGCTGCGCGCGCGAGCTGCACTTGGGCGAGATCAACA 241
 1041 CATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1100
 242 CATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 301
 1101 CGGATGCGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1133
 302 CGGATGCGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 334

RESULT 6
 AL545940 786 bp mRNA linear EST 16-FEB-2001
 LOCUS AL545940 LRT NFL006 PL2 Homo sapiens cDNA clone CS0D1023YK20 5

DEFINITION
 prime, mRNA sequence.

ACCESSION
 AL545940
 VERSION
 AL545940.1 GI:12878592
 KEYWORDS
 EST.

SOURCE

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 1 (bases 1 to 786)
 AUTHORS
 Li, W.B., Gruber, C., Jesses, J. and Polayes, D.
 TITLE
 Full-length cDNA libraries and normalization
 JOURNAL
 Unpublished (2001)
 COMMENT
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES

source

1. .786
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CS0D1023YK20"
 /clone_lib="LRT NFL006 PL2"
 /issue_type="Placenta"
 /note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA

was primed with a NotI-oligo (dt) primer. Five prime end

enriched, double-stranded cDNA was digested with Not I and

cloned into the Not I and Eco RV sites of the pCMVSPORT 6

vector. Library was normalized. Library was constructed by

life Technologies. Contact : Feng Liang Life Technologies,

a division of Invitrogen 9800 Medical Center Drive

Rockville, Maryland 20850, USA Fax : (1) 301 610 8371

Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com"

BASE COUNT 159 a 233 c 233 g 152 t 9 others
 ORIGIN

Query Match 27.8%; Score 329.4; DB 9; Length 786;
 Best Local Similarity 99.1%; Pred. No. 1.6e-31;
 Matches 330; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

801 AAAAGGCGCGGAGAGGGGTCTGCGACCGGCGCACTTGGCTGCTCCGCGCG 860
 1 AAAAGGCGCGGAGAGGGGTCTGCGACCGGCGCACTTGGCTGCTCCGCGCG 60
 861 CCACTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 920
 61 CCACTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
 921 CGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 980
 121 CGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
 981 GCGGAGCGCGAGCGCGCTGCGCGCGCGAGCTGCACTTGGGCGAGATCAACA 1040
 181 GCGGAGCGCGAGCGCGCTGCGCGCGCGAGCTGCACTTGGGCGAGATCAACA 240
 1041 CATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1100
 241 CATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300

1101 CGGATGCGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1133
 301 CGGATGCGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 333

RESULT 7
 B0072305 1197 bp mRNA linear EST 02-APR-2002
 LOCUS B0072305 AGENCOURT 6838866 NIH_MGC_122 Homo sapiens cDNA IMAGE:5761428

DEFINITION
 5', mRNA sequence.

ACCESSION
 B0072305
 VERSION
 B0072305.1 GI:19901351
 KEYWORDS
 EST.

SOURCE
 human.
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 1 (bases 1 to 1197)
 AUTHORS
 NIH-MGC http://mgi.mol.nih.gov/
 TITLE
 National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL
 Unpublished (1999)
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs.mol.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LIML at:
 http://image.llnl.gov
 Plate: IMAGE2810 row: a column: 13
 High quality sequence stop: 200.
 Location/Qualifiers

1. .1197
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5761428"

FEATURES

source

1. .1197
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5761428"

```
/clone_11b="NIH_MGC_122"
/lab host="DH10B"
/note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6;
Site 1: NotI; Site 2: EcoRV (destroyed); RNA source
anonymous pool of 24 week female lung, 16 week female
spleen, and 20-22 week male spleens. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.4 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Ivritrogen). Research Genetics tracking code 026. Note:
this is a NIH MGC Library."

BASE COUNT      274 a      310 c      493 g      120 t
ORIGIN

Query Match      27.4%; Score 325.6; DB 14; Length 1197;
Best Local Similarity 98.8%; Pred. No. 3.3e-31;
Matches 328; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY  802 AAAAGGCGCGGAAAGGCTCTGCAACCGGCACTTGAGCTCCGTCGGCGGCGC 861
Db  1  AAAAGCGCGGAAAGGCTCTGCAACCGGCACTTGAGCTCCGTCGGCGGCGC 60
QY  862 CACTTGCGCTGCTCCGTCGCGCGGCACTTGCGCTCCGTCGCGCGCGC 921
Db  61 CACTTGCGCTGCTCCGTCGCGCGGCACTTGCGCTCCGTCGCGCGCGC 120
QY  922 GGCATGCTGTGCGCGGCTGCGAGCTGCGCGCGGCTTGCGCGCGCGCA 981
Db  121 GGCATGCTGTGCGCGGCTGCGAGCTGCGCGCGGCTTGCGCGCGCGCA 180
QY  982 CGGAGCGCGGAGCGGCTGCGCGCGCAAGGAGTACTGCGGAGCAATCAAC 1041
Db  181 CGGAGCGCGGAGCGGCTGCGCGCGCAAGGAGTACTGCGGAGCAATCAAC 240
QY  1042 ATCTTCGCTGCGGCTGCGAGAGCAAGCAAGCGGAGCGGAGCTGTGATTC 1101
Db  241 ATCTTCGCTGCGGCTGCGAGAGCAAGCAAGCGGAGCGGAGCTGTGATTC 300
QY  1102 GGCATGCGGCGGCTACAGCTTGAAGGTGA 1133
Db  301 GGCATGCGGCGGCTACAGCTTGAAGGTGA 332

RESULT 8
BM917794      1622 bp      mRNA      linear      EST 12-MAR-2002
LOCUS      AGENCOURT 6614511 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:5485129
DEFINITION      5', mRNA sequence.
ACCESSION      BM917794
VERSION      BM917794.1 GI:19368173
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE      1 (bases 1 to 1622)
AUTHORS      NIH-MGC http://mgc.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: Dr. Daniel McVicar, DBS/NCI
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: L10CM2015 row: 1 column: 02
High quality sequence stop: 431.
Location/Qualifiers
1..1622
```

```
/organism="Homo sapiens"
/db xref="taxon:9606"
/clone="IMAGE:5485129"
/clone_11b="NIH_MGC_106"
/tissue type="natural killer cells, cell line"
/note="Organ: blood; Vector: pOTB1; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCGAGG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH MGC Library."

BASE COUNT      430 a      610 c      271 g      289 t      22 others
ORIGIN

Query Match      27.1%; Score 322.2; DB 14; Length 1622;
Best Local Similarity 99.1%; Pred. No. 6.5e-31;
Matches 324; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY  807 CGGCGGAAAGGAGTCTCTGCAACCGGCACTTGAGCTCCGTCGGCGGCGC 866
Db  1  CGGCGGAAAGGAGTCTCTGCAACCGGCACTTGAGCTCCGTCGGCGGCGC 60
QY  867 GGCCTGCTGCTCCGTCGCGCGGCACTTGCGCTCCGTCGCGCGGCGCAT 926
Db  61 GGCCTGCTGCTCCGTCGCGCGGCACTTGCGCTCCGTCGCGCGGCGCAT 120
QY  927 GCTGTGAGCGGAGCTGCGGCGCGGAGCTTGCGCGCGGAGCAAGAGCGAGA 986
Db  121 GCTGTGAGCGGAGCTGCGGAGCTTGCGCGCGGAGCTTGCGCGCGGAGCA 180
QY  987 CGCGGAGCGGCTGCGCGCGCAAGGAGTACTGCGGAGCAATCAACATCT 1046
Db  181 CGCGGAGCGGCTGCGCGCGCAAGGAGTACTGCGGAGCAATCAACATCT 240
QY  1047 CCGCTGCGGCTGCGAGAGCAAGCAAGCGGAGCGGAGCTGTGATTC 1106
Db  241 CCGCTGCGGCTGCGAGAGCAAGCAAGCGGAGCGGAGCTGTGATTC 300
QY  1107 GCGAGCGGCGCTACAGCTTGAAGGTGA 1133
Db  301 GCGAGCGGCGCTACAGCTTGAAGGTGA 327

RESULT 9
BQ434233      841 bp      mRNA      linear      EST 24-MAY-2002
LOCUS      AGENCOURT 7896761 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6159722
DEFINITION      5', mRNA sequence.
ACCESSION      BQ434233
VERSION      BQ434233.1 GI:2117309
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE      1 (bases 1 to 841)
AUTHORS      NIH-MGC http://mgc.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: ATCC/DCTD/DTF
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: L1AM13509 row: e column: 03
High quality sequence stop: 639.
```

FEATURES

SOURCE

Location/Qualifiers

1. 841
 /organism="Homo sapiens"
 /db xref="taxon:9606"
 /clone="IMAGR:6159722"
 /tissue_lib="NIH MGC 72"
 /tissue_type="melanotic melanoma"
 /lab host="DH10B (phage-resistant)"
 /note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
 Average insert size 2 kb. Library constructed by Life
 Technologies."

BASE COUNT 176 a 263 c 241 g 161 t
 ORIGIN

Query Match 27.0%; Score 320.4; DB 14; Length 841;
 Best Local Similarity 99.7%; Pred. No. 1.9e-30;
 Matches 321; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 812 GGAAGGGGCTCTGACACCGGCACTTGACCTGCTCCGCGCCGCACTTGAGCT 871
 DB 1 GGAAGGGGCTCTGACACCGGCACTTGACCTGCTCCGCGCCGCACTTGAGCT 60
 QY 872 GCTCCGCTCCGCGCGCGCACTTGCTGCTGCTCCGCGCGCGCGCATGCTG 931
 DB 61 GCTCCGCTCCGCGCGCGCACTTGCTGCTGCTCCGCGCGCGCGCATGCTG 120
 QY 932 TGGCGGCTCGAGACTGCGCGCGCGCTTGCCTCCCGCGCGCAAGAGCGGAGCGCG 991
 DB 121 TGGCGGCTCGAGACTGCGCGCGCGCTTGCCTCCCGCGCGCAAGAGCGGAGCGCG 180
 QY 992 AGCCGCGTCCGCGCGCGCAAGAGACTGCAATCTGCGCGCGCAATCTCCGCT 1051
 DB 181 AGCCGCGTCCGCGCGCGCAAGAGACTGCAATCTGCGCGCGCAATCTCCGCT 240
 QY 1052 GGGGCTGAG 1111
 DB 241 GGGGCTGAG 300
 QY 1112 CGCGTACAGCCTGAGAGTGA 1133
 DB 301 CGCGTACAGCCTGAGAGATGA 322

RESULT 10

BM66216

LOCUS BM66216 985 bp mRNA linear EST 05-FEB-2002
 DEFINITION AGENCOURT_6456956 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:5576962
 5', mRNA sequence.

ACCESSION

BM66216
 BM66216.1 GI:18515258

KEYWORDS

EST.

SOURCE

ORGANISM

human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE
 1 (bases 1 to 985)
 NIH-MGC http://mgi.nci.nih.gov/.
 TITLE
 National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL
 Unpublished (1999)
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: ggapbs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LHAM12329 row: k column: 11
 High quality sequence stop: 522.
 Location/Qualifiers

FEATURES

SOURCE

1. 985

/organism="Homo sapiens"
 /db xref="taxon:9606"
 /clone="IMAGR:5576962"
 /tissue_lib="NIH MGC 92"
 /tissue_type="embryonal carcinoma, cell line"
 /lab host="DH10B (phage-resistant)"
 /note="Organ: testis; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Oligo-dt primed.
 Average insert size 2.5 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH MGC library."

BASE COUNT 194 a 323 c 264 g 204 t
 ORIGIN

Query Match 26.7%; Score 317.4; DB 13; Length 985;
 Best Local Similarity 98.2%; Pred. No. 3.9e-30;
 Matches 321; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 807 CGCGGGAAGGGGCTGCGACCGGCACTTGAGCTGCTCCGCGCGCGCACTT 866
 DB 4 CAAGGCTCCGCGGCTGCGACCGGCACTTGAGCTGCTCCGCGCGCGCACTT 63
 QY 867 GGCCTGCTCCGCGCGCGCACTTGCTGCTGCTCCGCGCGCGCGCAT 926
 DB 64 GGCCTGCTCCGCGCGCGCACTTGCTGCTGCTCCGCGCGCGCGCAT 123
 QY 927 GCTGTGCGCGCTCGAGAGTGGCGCGCGCTTGCCTCCCGCGCGCAAGAGCGGGA 986
 DB 124 GCTGTGCGCGCTCGAGAGTGGCGCGCGCTTGCCTCCCGCGCGCAAGAGCGGGA 183
 QY 987 CGCGGAGCGCGCTCGCGCGCGAGAGTGAATCTGAGTGGGCGAGATCCACATCTT 1046
 DB 184 CGCGGAGCGCGCTCGCGCGCGAGAGTGAATCTGAGTGGGCGAGATCCACATCTT 243
 QY 1047 CGCGTCCGCGCTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1106
 DB 244 CGCGTCCGCGCTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 303
 QY 1107 GCAAGGCGGCTACAGCCTGAGAGTGA 1133
 DB 304 GCAAGGCGGCTACAGCCTGAGAGATGA 330

RESULT 11

BM800322

LOCUS BM800322 1018 bp mRNA linear EST 05-MAR-2002
 DEFINITION AGENCOURT_6420080 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:5584389
 5', mRNA sequence.

ACCESSION

BM800322
 BM800322.1 GI:19117445

KEYWORDS

EST.

SOURCE

ORGANISM

human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE
 1 (bases 1 to 1018)
 NIH-MGC http://mgi.nci.nih.gov/.
 TITLE
 National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL
 Unpublished (1999)
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: ggapbs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LHAM12348 row: p column: 22
 High quality sequence stop: 583.
 Location/Qualifiers

FEATURES

SOURCE

1. 1018

/organism="Homo sapiens"


```
/db_xref="taxon:9606"
/clone="IMAGE:5584389"
/clone_1lb="NIH_MGC_92"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/Note="Organ: testis; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

BASE COUNT      217 a      322 c      273 g      206 t
ORIGIN

Query Match      26.7%; Score 317.2; DB 14; Length 1018;
Best Local Similarity 99.1%; Pred. No. 4e-30; Indels 0; Gaps 0;
Matches 319; Conservative 0; Mismatches 3;

QY 812 GGAAGGGGCTCTGCGACCGGCACTTGAGCTGCGTCCGCGCGCGCACTTGAGCT 871
Db 2 GGAAGGGGCTCTGCGACCGGCACTTGAGCTGCGTCCGCGCGCGCACTTGAGCT 61

QY 872 GCGTCCGCTCCGCGCGCGCACTTGAGCTGCGTCCGCGCGCGCACTTGAGCT 931
Db 62 GCGTCCGCTCCGCGCGCGCACTTGAGCTGCGTCCGCGCGCGCACTTGAGCT 121

QY 932 TGGCCGCTCCGCGCGCGCACTTGAGCTGCGTCCGCGCGCGCACTTGAGCT 991
Db 122 TGGCCGCTCCGCGCGCGCACTTGAGCTGCGTCCGCGCGCGCACTTGAGCT 181

QY 992 AGCCGCGCTCCGCGCGCGCACTTGAGCTGCGTCCGCGCGCGCACTTGAGCT 1051
Db 182 AGCCGCGCTCCGCGCGCGCACTTGAGCTGCGTCCGCGCGCGCACTTGAGCT 241

QY 1052 GCGGCGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1111
Db 242 GCGGCGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 301

QY 1112 CGCGCTACAGCTTGAGAGTGGA 1133
Db 302 CGCGCTACAGCTTGAGAGTGGA 323

RESULT 12
BQ685120 1021 bp mRNA linear EST 15-JUL-2002
LOCUS BQ685120
DEFINITION AGENCOURT 8344472 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6251301
5' mRNA sequence.
ACCESSION BQ685120
VERSION BQ685120.1 GI:21810436
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 1021)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: L10M2395 row: 1 column: 22
High quality sequence start: 12
High quality sequence stop: 409.
Location/Qualifiers
1..1021
/organism="Homo sapiens"

FEATURES
source

/db_xref="taxon:9606"
/clone="IMAGE:6251301"
/clone_1lb="NIH_MGC_110"
/tissue_type="ductal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/Note="Organ: pancreas; Vector: pORF7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming;
directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GCGACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

BASE COUNT      202 a      335 c      289 g      194 t      1 others
ORIGIN

Query Match      26.3%; Score 312.2; DB 14; Length 1021;
Best Local Similarity 99.1%; Pred. No. 1.6e-29; Indels 0; Gaps 0;
Matches 314; Conservative 0; Mismatches 3;

QY 817 GGGTCTGCGACCGGCGCACTTGAGCTGCGTCCGCGCGCGCACTTGAGCT 876
Db 7 GGGTCTGCGACCGGCGCACTTGAGCTGCGTCCGCGCGCGCACTTGAGCT 66

QY 877 CGTCCGCGCGCGCACTTGAGCTGCGTCCGCGCGCGCACTTGAGCT 936
Db 67 CGTCCGCGCGCGCACTTGAGCTGCGTCCGCGCGCGCACTTGAGCT 126

QY 937 GGGTCCGCGCGCGCACTTGAGCTGCGTCCGCGCGCGCACTTGAGCT 996
Db 127 GGGTCCGCGCGCGCACTTGAGCTGCGTCCGCGCGCGCACTTGAGCT 186

QY 997 CGTCCGCGCGCGCACTTGAGCTGCGTCCGCGCGCGCACTTGAGCT 1056
Db 187 CGTCCGCGCGCGCACTTGAGCTGCGTCCGCGCGCGCACTTGAGCT 246

QY 1057 GTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1116
Db 247 GTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 306

QY 1117 TACAGCTTGAGAGTGGA 1133
Db 307 TACAGCTTGAGAGTGGA 323

RESULT 13
AL541351 668 bp mRNA linear EST 16-FEB-2001
LOCUS AL541351
DEFINITION AL541351 LTI FL002_P11 Homo sapiens cDNA clone CSDBE006YA08 5 prime
mRNA sequence.
ACCESSION AL541351
VERSION AL541351.1 GI:12872339
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 668)
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLES Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
1..668
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CSDBE006YA08"
/clone_1lb="LTI FL002_P11"
/lab_host="DH10B"
/Note="Organ: placenta; Vector: pCMVSPORT 6; 1st strand
```

```

/clone lib="EPI_NFL006_P12"
/issue type="placements"
/notes="vector: pGEMT01 6: site 1: NotI; left strand cDNA
was primed, with a NotI-oligo (34) primer. Five prime end
enriched, double-stranded cDNA was digested with NotI and
cloned into the NotI and Eco RV sites of the pGEMT01 6
vector. Library was normalized. Library was constructed by

```

```

/Organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="S13KMS5-14-F09"
/clone_lib="S13KMS5"
/tissue_type="myeloma"
/cell_line="KMS-5"
/lab_host="Top10F"

```

/note="Vector: PCNS; Site_1: BcoRI; Site_2: NotI; The poly (A) + RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then decapped with tobacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA-RNA linker including BcoRI site by treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dT-selected mRNA by priming with dT-tailed vector. The dT-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of BcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

BASE COUNT 89 a 173 c 163 g 98 t
ORIGIN

Query Match 25.1%; Score 298.4; DB 14; Length 523;
Best Local Similarity 99.7%; Pred. No. 1.4e-27;
Matches 299; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 834 CACTTGGCTGCTCCGTCGCGCGCGCACTTGGCTGCTCCGTCGCGCGCGCACT 893
|||
DB 2 CACTTGGCTGCTCCGTCGCGCGCGCACTTGGCTGCTCCGTCGCGCGCGCACT 61
QY 894 TCGCTGCTCCGTCGCGCGCGCGCGCACTGCTGCTGCTGCTGCTGCTGCTGCT 953
|||
DB 62 TCGCTGCTCCGTCGCGCGCGCGCGCGCACTGCTGCTGCTGCTGCTGCTGCTGCT 121
QY 954 CCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1013
|||
DB 122 CCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 181
QY 1014 GCTGAGTACCTGGGAGATCAACATCTCTGCTGCTGCTGCTGCTGCTGCTGCT 1073
|||
DB 182 GCTGAGTACCTGGGAGATCAACATCTCTGCTGCTGCTGCTGCTGCTGCTGCT 241
QY 1074 CACGGGACCGGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1133
|||
DB 242 CACGGGACCGGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 301

Search completed: November 28, 2002, 22:40:23
Job time : 1582.2 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 28, 2002, 19:48:24 ; Search time 92.4142 Seconds

(without alignments)
8693.659 Million cell updates/sec

Title: US-09-963-333-7

Perfect score: 1187
Sequence: 1 gatcgccacatcgactcca.....gagggagcgcgctg999a 1187Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 344316 seqs, 338423730 residues

Total number of hits satisfying chosen parameters: 688632

Minimum DB seq length: 0
Maximum DB seq length: 200000000Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Pending Patents NA New:*
1: /cgn2_6/ptodata/2/pna/US06_NMW_COMB.seq:*
2: /cgn2_6/ptodata/2/pna/US06_NMW_COMB.seq:*
3: /cgn2_6/ptodata/2/pna/US07_NMW_COMB.seq:*
4: /cgn2_6/ptodata/2/pna/US08_NMW_COMB.seq:*
5: /cgn2_6/ptodata/2/pna/US09_NMW_COMB.seq:*
6: /cgn2_6/ptodata/2/pna/US10_NMW_COMB.seq:*
7: /cgn2_6/ptodata/2/pna/US60_NMW_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1178	99.2	1187	5	US-09-658-659B-7	Sequence 7, Appli
2	1167	98.3	18597	5	US-09-658-659B-8	Sequence 8, Appli
3	770.6	64.9	1186	6	US-10-257-166-112	Sequence 172, Appl
4	710	59.8	1186	6	US-10-257-166-171	Sequence 171, Appl
5	301	25.4	1942	5	US-09-724-676-34523	Sequence 34523, A
6	301	25.4	1942	5	US-09-724-676A-34523	Sequence 34523, A
7	301	25.4	2042	5	US-09-724-676-34515	Sequence 34515, A
8	301	25.4	2042	5	US-09-724-676A-34515	Sequence 34515, A
9	301	25.4	2391	5	US-09-724-676-34597	Sequence 34597, A
10	301	25.4	2391	5	US-09-724-676A-34597	Sequence 34597, A
11	301	25.4	2411	5	US-09-724-676-34535	Sequence 34535, A
12	301	25.4	2411	5	US-09-724-676A-34535	Sequence 34535, A
13	301	25.4	2491	5	US-09-724-676-34589	Sequence 34589, A
14	301	25.4	2491	5	US-09-724-676A-34589	Sequence 34589, A
15	301	25.4	2527	5	US-09-724-676-34548	Sequence 34548, A
16	301	25.4	2527	5	US-09-724-676A-34548	Sequence 34548, A
17	301	25.4	3023	5	US-09-724-676-34565	Sequence 34565, A
18	301	25.4	3023	5	US-09-724-676A-34565	Sequence 34565, A
19	301	25.4	3123	5	US-09-724-676-34557	Sequence 34557, A
20	301	25.4	3123	5	US-09-724-676A-34557	Sequence 34557, A
21	301	25.4	3139	5	US-09-724-676-34581	Sequence 34581, A
22	301	25.4	3139	5	US-09-724-676A-34581	Sequence 34581, A
23	301	25.4	3239	5	US-09-724-676-34573	Sequence 34573, A
24	301	25.4	3239	5	US-09-724-676A-34573	Sequence 34573, A
25	299.4	25.4	1536	5	US-09-658-659B-6	Sequence 6, Appli
26	288.8	24.3	3298	6	US-10-240-965-200	Sequence 200, Appl

27	204	17.2	304905	1	PCT-US02-32700-1	Sequence 1, Appli
28	204	17.2	304905	6	US-10-271-416-1	Sequence 1, Appli
29	199	16.8	135259	6	US-10-240-425-1585	Sequence 1585, Ap
30	198.4	16.7	207433	6	US-10-277-216-5	Sequence 5, Appli
31	193.4	16.3	162025	6	US-10-272-665-35	Sequence 35, Appli
32	193.4	16.3	162025	6	US-10-273-321-35	Sequence 35, Appli
33	193.4	16.3	162025	6	US-10-272-756-35	Sequence 35, Appli
34	193.4	16.3	162025	6	US-10-272-756-36	Sequence 36, Appli
35	193.4	16.3	162025	6	US-10-272-756-36	Sequence 36, Appli
36	193.4	16.3	162025	6	US-10-273-228-35	Sequence 35, Appli
37	193.4	16.3	162025	6	US-10-273-228-36	Sequence 36, Appli
38	192.6	16.2	202100	1	PCT-US02-34679-484	Sequence 484, App
39	192.6	16.2	202100	6	US-10-282-174-484	Sequence 484, App
40	191.2	16.1	36534	6	US-10-240-425-1461	Sequence 1461, Ap
41	191.2	16.1	36534	6	US-10-240-425-1461	Sequence 1461, Ap
42	191.2	16.1	145831	6	US-10-240-425-1470	Sequence 1470, App
43	191.2	16.1	100671	6	US-10-240-425-1470	Sequence 1470, App
44	189.2	15.9	396	5	US-09-513-999C-28575	Sequence 28575, A
45	189.2	15.9	201143	6	US-10-240-425-1099	Sequence 1099, Ap

ALIGNMENTS

RESULT 1
US-09-658-659B-7
Sequence 7, Application US/09658659B
GENERAL INFORMATION:
APPLICANT: Stanton, Jr., Vincent P.
TITLE OF INVENTION: GENE SEQUENCE VARIANCE IN GENES RELATED
TO FOLATE METABOLISM HAVING UTILITY IN DETERMINING THE
TITLE OF INVENTION: TREATMENT OF DISEASE
FILE REFERENCE: 11926-015001
CURRENT APPLICATION NUMBER: US/09/658,659B
CURRENT FILING DATE: 2000-09-08
PRIORITY APPLICATION NUMBER: 09/556,033
PRIORITY FILING DATE: 2000-06-15
PRIORITY APPLICATION NUMBER: 09/357,743
PRIORITY FILING DATE: 1999-07-20
PRIORITY APPLICATION NUMBER: 09/357,024
PRIORITY FILING DATE: 1999-07-19
PRIORITY APPLICATION NUMBER: 60/093,484
PRIORITY FILING DATE: 1998-07-20
NUMBER OF SEQ ID NOS: 16
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7
LENGTH: 1187
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: 276_321, 534, 656
OTHER INFORMATION: n = c or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: 452_640
OTHER INFORMATION: n = a or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: 492, 625
OTHER INFORMATION: n = c or a
OTHER INFORMATION: nucleotide in position 458 is c, or absent
US-09-658-659B-7
Query Match 99.2%, Score 1178, DB 5, Length 1187;
Best Local Similarity 100.0%; Pred. No. 4.8e-151;
Matches 1187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 GATCGGCACATCGACCTCGAGGAGCGGAGCTGCTCAAAAAA 60
|||||

Db 646 CCGCAAAAAAGTAAATCTACTACTTAATAAAAAAGCGATGCGCAAAAGAT 587
QY 602 TCCCAAAAGGCGAGTCTTCCGCGCACCGCACCTGAGTCCAGGTTCCGGGTTTCTA 661
Db 586 TCCCAAAAGGCGAGTCTTCCGCGCACCGCACCTGAGTCCAGGTTCCGGGTTTCTA 527
QY 662 AGACTCTAGCTGTGGCCCTGGGCTCGTTCTGTCGACACCGGTGGCTCTCGCTTCC 721
Db 526 AAGCTCTCACTATTAACCTTAACCTCGTTTCTATACACACCGGTAACTCTCACTTCC 467
QY 722 CCTTGGGCGACGCTCTTGAAGGCGGCGCGCGGACCCCGCGAAGGAAAGGCGG 781
Db 466 CCTTAACGCGACGCTCTTGAAGGCGGCGCGGACCCCGCGAAGGAAAGGCGG 407
QY 782 AAGCGGGAAGCGCGCGGGAAGGAAAGGCGGGAAGGCGGCGGCGGCGGCGGCGG 841
Db 406 AAGCGGGAAGCGCGCGGGAAGGAAAGGCGGGAAGGCGGCGGCGGCGGCGGCGG 347
QY 842 CTGCGCTCGGTCGCGCGCGCGCACCTTGGCTCGCTCGCTCGCGCGCGCGCACCTTGG 901
Db 346 CTACCTTCGTTCCGCGCGCGCGCACCTTACCTTCCGTCGCGCGCGCGCACCTTGG 287
QY 902 CTGCGCTCGGTCGCGCGCGCGCACCTTGGCTCGCTCGCTCGCGCGCGCGCACCTTGG 961
Db 286 CTGCGCTCGGTCGCGCGCGCGCACCTTACCTTACCGGCTGGAACGCGCGCGCACCTT 227
QY 962 TCCGCGCGCGCGCGCGCGCGCGCGCGCGCTGTCGCGCGCGCGCGCGCGCGCGCT 1021
Db 226 TACCG 167
QY 1022 ACCTGGGCGCATTCACACACATCTCGCTGCGGCGGTCGAGGAGGACGCGCGCGGCA 1081
Db 166 ACCTGAAACCAATCAACACATCTCGCTGAGAGCTGAAAGGAGGAGGAGGAGGAG 107
QY 1082 CCGGACCGCTGCGGTAATTCGCGCATGCGAGCGCGCTGACGCTGAGAGTGAAGCGCGG 1141
Db 106 CCGACACCTTATCGATATTCGACATACAAACGCGCTACACCTTAAAGGAGGCGGCGG 47
QY 1142 GCCCTGGCGGAGCGGTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1182
Db 46 ACCCTTACGAAACGATTAACGAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 6

RESULT 4
US-10-257-166-171
Sequence 171, Application US/10257166
GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: PIEPENBROCK, Christian
APPLICANT: BERLIN, Kurt
TITLE OF INVENTION: Method and Nucleic Acids for Analysing the Methylation of
TITLE OF INVENTION: Genes Implicated in Pharmacogenomics
FILE REFERENCE: 5013.1011
CURRENT APPLICATION NUMBER: US/10/257.166
CURRENT FILING DATE: 2002-10-07
PRIOR APPLICATION NUMBER: PCT/EP01/07470
DE 10032529.7
DE 10043826.1
PRIOR FILING DATE: 2001-06-29
2000-06-30
2000-09-01
NUMBER OF SEQ ID NOS: 178
SEQ ID NO 171
LENGTH: 1186
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-257-166-171

Query Match 59.8%; Score 710; DB 6; Length 1186;
Best Local Similarity 75.1%; Pctd. No. 4, 5e-88;
Matches 892; Conservative 0; Mismatches 294; Indels 1; Gaps 1;

QY 1 GATGCGCGCATGCGACTCCAGCGCTGGGTGAGAGCGGAGCTGTGTCAAAAAA 60
Db 1 GATGCGCGCATGCGACTCCAGCGCTGGGTGAGAGCGGAGCTGTGTCAAAAAA 60
QY 61 AAAAAAGCCGCGAGGCTCAACAAAAAAGCTCGGAAAGCCCTGGCGGCTCTTTT 120
Db 61 AAAAAAGCCGCGAGGCTCAACAAAAAAGCTCGGAAAGCCCTGGCGGCTCTTTT 120
QY 121 TTTTCTTTTCTTTTCTTTTGGGACAGCTGTGCTGTGCGCCGAGGCGGAGGAGG 180
Db 121 TTTTCTTTTCTTTTCTTTTGGGACAGCTGTGCTGTGCGCCGAGGCGGAGGAGG 180
QY 181 CGGACTTGGCTGCTGCAACCTCGCGCTCCAGGTCAGAGCAATCTTCTGCTCAGCC 240
Db 181 CGGACTTGGCTGCTGCAACCTCGCGCTCCAGGTCAGAGCAATCTTCTGCTCAGCC 240
QY 241 TCCCAAGTACGCAACGCGCGCTAATTTTGTATTTTGTATGAGAGCGGCGGTTT 300
Db 241 TTTTATGATGATTAATTAAGCTTATGATTTTGTATTTTATGAGAGCGGCGGTTT 300
QY 301 CATGTGTCAGGCTGCTGCTGCACTCTGACCTGAGGTATCCACCCGCTCGGCCCC 360
Db 301 CATGTGTCAGGCTGCTGCTGCACTCTGACCTGAGGTATCCACCCGCTCGGCCCC 360
QY 361 TATGTTGTTAGGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 420
Db 361 TATGTTGTTAGGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 420
QY 421 AGTAAAGAGCTGATTAACCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
Db 421 AGTAAAGAGCTGATTAACCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
QY 481 AATGCAATCTGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATG 540
Db 481 AATGCAATCTGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATG 540
QY 541 AATGCAATCTGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATG 599
Db 541 AATGCAATCTGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATG 599
QY 599 AATGCAATCTGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATG 660
Db 599 AATGCAATCTGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATG 660
QY 660 TTTTCAAGGCGGAGTCTTTTCAAGGCGGAGTCTTTTCAAGGCGGAGTCTTTT 719
Db 660 TTTTCAAGGCGGAGTCTTTTCAAGGCGGAGTCTTTTCAAGGCGGAGTCTTTT 719
QY 719 TTTTCAAGGCGGAGTCTTTTCAAGGCGGAGTCTTTTCAAGGCGGAGTCTTTT 780
Db 719 TTTTCAAGGCGGAGTCTTTTCAAGGCGGAGTCTTTTCAAGGCGGAGTCTTTT 780
QY 780 TTTTCAAGGCGGAGTCTTTTCAAGGCGGAGTCTTTTCAAGGCGGAGTCTTTT 840
Db 780 TTTTCAAGGCGGAGTCTTTTCAAGGCGGAGTCTTTTCAAGGCGGAGTCTTTT 840
QY 840 TTTTCAAGGCGGAGTCTTTTCAAGGCGGAGTCTTTTCAAGGCGGAGTCTTTT 899
Db 840 TTTTCAAGGCGGAGTCTTTTCAAGGCGGAGTCTTTTCAAGGCGGAGTCTTTT 899
QY 899 TTTTCAAGGCGGAGTCTTTTCAAGGCGGAGTCTTTTCAAGGCGGAGTCTTTT 960
Db 899 TTTTCAAGGCGGAGTCTTTTCAAGGCGGAGTCTTTTCAAGGCGGAGTCTTTT 960
QY 960 TTTTCAAGGCGGAGTCTTTTCAAGGCGGAGTCTTTTCAAGGCGGAGTCTTTT 1019
Db 960 TTTTCAAGGCGGAGTCTTTTCAAGGCGGAGTCTTTTCAAGGCGGAGTCTTTT 1019
QY 1019 TTTTCAAGGCGGAGTCTTTTCAAGGCGGAGTCTTTTCAAGGCGGAGTCTTTT 1080
Db 1019 TTTTCAAGGCGGAGTCTTTTCAAGGCGGAGTCTTTTCAAGGCGGAGTCTTTT 1080
QY 1080 TTTTCAAGGCGGAGTCTTTTCAAGGCGGAGTCTTTTCAAGGCGGAGTCTTTT 1080
Db 1080 TTTTCAAGGCGGAGTCTTTTCAAGGCGGAGTCTTTTCAAGGCGGAGTCTTTT 1080

CURRENT APPLICATION NUMBER: US/09/724,676
CURRENT FILING DATE: 2000-11-28
NUMBER OF SEQ ID NOS: 97222
SOFTWARE: PatentIn version 3.2
SEQ ID NO 34535
LENGTH: 2411
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (41)-(41)
OTHER INFORMATION: n is a,c,g, or t
US-09-724-676-34535

Query Match 25.4%; Score 301; DB 5; Length 2411;
Best Local Similarity 100.0%; Pred. No. 4.5e-33;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 833 CCACTTGACCTCCCTCCGCGCCGACATGAGCTGCTCCGCGCCGCGCAC 892
Db 2398 CCACTTGACCTCCCTCCGCGCCGACATGAGCTGCTCCGCGCCGCGCAC 2339
Qy 893 TTGCTGCTCCCTCCGCGCCGACATGAGCTGCTCCGCGCCGCGCAC 952
Db 2338 TTGCTGCTCCCTCCGCGCCGACATGAGCTGCTCCGCGCCGCGCAC 2279
Qy 953 GCCGCGCTTGGCCCGCGCGACAGAGCGGAGCGCGCGCGCGCGCGCG 1012
Db 2278 GCCGCGCTTGGCCCGCGCGACAGAGCGGAGCGCGCGCGCGCGCGCG 2219
Qy 1013 AGCTGAGTACCTGGGGGAGATCCAAACATCTCTCGCTGCGGCGGTAGAGAGAGAGACC 1072
Db 2218 AGCTGAGTACCTGGGGGAGATCCAAACATCTCTCGCTGCGGCGGTAGAGAGAGAGACC 2159
Qy 1073 GCACGGGACCGGACCTCTGCGGTATTCGAGCGCGGTACAGCTGAGAGGTG 1132
Db 2158 GCACGGGACCGGACCTCTGCGGTATTCGAGCGCGGTACAGCTGAGAGGTG 2099
Qy 1133 A 1133
Db 2098 A 2098

RESULT 12
US-09-724-676A-34535/c
Sequence 34535, Application US/09724676A
GENERAL INFORMATION:
APPLICANT: Compugen LTD
TITLE OF INVENTION: Variants of alternative splicing
FILE REFERENCE: 129181.4 Compugen
CURRENT APPLICATION NUMBER: US/09/724,676A
CURRENT FILING DATE: 2000-11-28
NUMBER OF SEQ ID NOS: 97222
SOFTWARE: PatentIn version 3.2
SEQ ID NO 34535
LENGTH: 2411
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (41)-(41)
OTHER INFORMATION: n is a,c,g, or t
US-09-724-676A-34535

Query Match 25.4%; Score 301; DB 5; Length 2411;
Best Local Similarity 100.0%; Pred. No. 4.5e-33;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 833 CCACTTGACCTCCCTCCGCGCCGACATGAGCTGCTCCGCGCCGCGCAC 892
Db 2398 CCACTTGACCTCCCTCCGCGCCGACATGAGCTGCTCCGCGCCGCGCAC 2339
Qy 893 TTGCTGCTCCCTCCGCGCCGACATGAGCTGCTCCGCGCCGCGCAC 952

Db 2338 TTGCTGCTCCCTCCGCGCCGACATGAGCTGCTCCGCGCCGCGCAC 2279
Qy 953 GCCGCGCTTGGCCCGCGCGACAGAGCGGAGCGCGCGCGCGCGCGCG 1012
Db 2278 GCCGCGCTTGGCCCGCGCGACAGAGCGGAGCGCGCGCGCGCGCGCG 2219
Qy 1013 AGCTGAGTACCTGGGGGAGATCCAAACATCTCTCGCTGCGGCGGTAGAGAGAGACC 1072
Db 2218 AGCTGAGTACCTGGGGGAGATCCAAACATCTCTCGCTGCGGCGGTAGAGAGAGACC 2159
Qy 1073 GCACGGGACCGGACCTCTGCGGTATTCGAGCGCGGTACAGCTGAGAGGTG 1132
Db 2158 GCACGGGACCGGACCTCTGCGGTATTCGAGCGCGGTACAGCTGAGAGGTG 2099
Qy 1133 A 1133
Db 2098 A 2098

RESULT 13
US-09-724-676-34589/c
Sequence 34589, Application US/09724676
GENERAL INFORMATION:
APPLICANT: Compugen LTD
TITLE OF INVENTION: Variants of alternative splicing
FILE REFERENCE: 129181.4 Compugen
CURRENT APPLICATION NUMBER: US/09/724,676
CURRENT FILING DATE: 2000-11-28
NUMBER OF SEQ ID NOS: 97222
SOFTWARE: PatentIn version 3.2
SEQ ID NO 34589
LENGTH: 2491
TYPE: DNA
ORGANISM: Homo sapiens
US-09-724-676-34589

Query Match 25.4%; Score 301; DB 5; Length 2491;
Best Local Similarity 100.0%; Pred. No. 4.5e-33;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 833 CCACTTGACCTCCCTCCGCGCCGACATGAGCTGCTCCGCGCCGCGCAC 892
Db 2478 CCACTTGACCTCCCTCCGCGCCGACATGAGCTGCTCCGCGCCGCGCAC 2419
Qy 893 TTGCTGCTCCCTCCGCGCCGACATGAGCTGCTCCGCGCCGCGCAC 952
Db 2418 TTGCTGCTCCCTCCGCGCCGACATGAGCTGCTCCGCGCCGCGCAC 2359
Qy 953 GCCGCGCTTGGCCCGCGCGACAGAGCGGAGCGCGCGCGCGCGCGCG 1012
Db 2358 GCCGCGCTTGGCCCGCGCGACAGAGCGGAGCGCGCGCGCGCGCGCG 2299
Qy 1013 AGCTGAGTACCTGGGGGAGATCCAAACATCTCTCGCTGCGGCGGTAGAGAGAGACC 1072
Db 2298 AGCTGAGTACCTGGGGGAGATCCAAACATCTCTCGCTGCGGCGGTAGAGAGAGACC 2239
Qy 1073 GCACGGGACCGGACCTCTGCGGTATTCGAGCGCGGTACAGCTGAGAGGTG 1132
Db 2238 GCACGGGACCGGACCTCTGCGGTATTCGAGCGCGGTACAGCTGAGAGGTG 2179
Qy 1133 A 1133
Db 2178 A 2178

RESULT 14
US-09-724-676A-34589/c
Sequence 34589, Application US/09724676A
GENERAL INFORMATION:
APPLICANT: Compugen LTD
TITLE OF INVENTION: Variants of alternative splicing
FILE REFERENCE: 129181.4 Compugen

;; CURRENT APPLICATION NUMBER: US/09/724,676A
;; CURRENT FILING DATE: 2000-11-28
;; NUMBER OF SEQ ID NOS: 97222
;; SOFTWARE: PatentIn version 3.2
;; SEQ ID NO 34589
;; LENGTH: 2491
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-724-676A-34589

Query Match 25.4%; Score 301; DB 5; Length 2491;
Best Local Similarity 100.0%; Pred. No. 4.Se-33;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 833 CCACTTGAGCTGCTCCGCTCCGCGCCGCGCACTTGAGCTGCTCCGCTCCGCGCGCCAC 892
DB 2478 CCACTTGAGCTGCTCCGCTCCGCGCCGCGCACTTGAGCTGCTCCGCTCCGCGCGCCAC 2419
QY 893 TTGCGCTGCTCCGCTCCG 952
DB 2418 TTGCGCTGCTCCGCTCCG 2359
QY 953 GCCGCGCTTGCCTCCG 1012
DB 2358 GCCGCGCTTGCCTCCG 2299
QY 1013 AGCTGCACTGCTGCG 1072
DB 2298 AGCTGCACTGCTGCG 2239
QY 1073 GACGCGGACCG 1132
DB 2238 GACGCGGACCG 2179
QY 1133 A 1133
DB 2178 A 2178

RESULT 15

US-09-724-676-34548/C
;; Sequence 34548, Application US/09724676
;; GENERAL INFORMATION:
;; APPLICANT: Compugen LTD
;; TITLE OF INVENTION: Variants of alternative splicing
;; FILE REFERENCE: 129181.4 Compugen
;; CURRENT APPLICATION NUMBER: US/09/724,676
;; CURRENT FILING DATE: 2000-11-28
;; NUMBER OF SEQ ID NOS: 97222
;; SOFTWARE: PatentIn version 3.2
;; SEQ ID NO 34548
;; LENGTH: 2527
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (41)..(41)
;; OTHER INFORMATION: n is a,c,g, or t
US-09-724-676-34548

Query Match 25.4%; Score 301; DB 5; Length 2527;
Best Local Similarity 100.0%; Pred. No. 4.Se-33;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 833 CCACTTGAGCTGCTCCGCTCCGCGCCGCGCACTTGAGCTGCTCCGCTCCGCGCGCCAC 892
DB 2514 CCACTTGAGCTGCTCCGCTCCGCGCCGCGCACTTGAGCTGCTCCGCTCCGCGCGCCAC 2455
QY 893 TTGCGCTGCTCCGCTCCG 952
DB 2454 TTGCGCTGCTCCGCTCCG 2395
QY 953 GCCGCGCTTGCCTCCG 1012

DB 2394 GCCGCGCTTGCCTCCG 2335
QY 1013 AGCTGCACTGCTGCG 1072
DB 2334 AGCTGCACTGCTGCG 2275
QY 1073 GACGCGGACCG 1132
DB 2274 GACGCGGACCG 2215
QY 1133 A 1133
DB 2214 A 2214

Search completed: November 29, 2002, 00:20:29
Job time : 115.914 secs

QY 1081 GTCTTTAGGGGTTGGCGTGAATCCGAGGTAAAGTTCTTTTGTCTTAAAGANAAG 1140
DB 1081 GTCTTTAGGGGTTGGCGTGAATCCGAGGTAAAGTTCTTTTGTCTTAAAGANAAG 1140
QY 1141 AACTAGGTCAAAAATCTGTCCGTGACCTTATGATTAATTTTAAAGANTTCCACT 1200
DB 1141 AACTAGGTCAAAAATCTGTCCGTGACCTTATGATTAATTTTAAAGANTTCCACT 1200
QY 1201 GCGAATGTAATCTGTCAGTCTTTCATATAAAGGCTTTGAGTAACTCACTAGAG 1260
DB 1201 GCGAATGTAATCTGTCAGTCTTTCATATAAAGGCTTTGAGTAACTCACTAGAG 1260
QY 1261 GTATCTGACAAATGCTGAGTTATGAACAAAGTGAAGAAAGTAATGTCCTTAG 1320
DB 1261 GTATCTGACAAATGCTGAGTTATGAACAAAGTGAAGAAAGTAATGTCCTTAG 1320
QY 1321 CAATAACATGATGATGCAATTCATCCAGTACTTAAAGAGGTTGATTTTCAC 1380
DB 1321 CAATAACATGATGATGCAATTCATCCAGTACTTAAAGAGGTTGATTTTCAC 1380
QY 1381 AAGCTATTTTGAATATTTTGAATATTTTGAATATTTTGAATATTTTGAATAT 1440
DB 1381 AAGCTATTTTGAATATTTTGAATATTTTGAATATTTTGAATATTTTGAATAT 1440
QY 1441 CTGAGGAGCTGAGTAAACACCATGATCATGATGTAAGTGTGTATGAACTTTA 1496
DB 1441 CTGAGGAGCTGAGTAAACACCATGATCATGATGTAAGTGTGTATGAACTTTA 1496

RESULT 7
US-09-442-384A-792
; Sequence 792, Application US/0942384A
; GENERAL INFORMATION:
; APPLICANT: Chenchik, Alex
; APPLICANT: Lukashiev, Matvey
; TITLE OF INVENTION: Hematology/Immunology Array
; FILE REFERENCE: CLON-006CIP15
; CURRENT APPLICATION NUMBER: US/09/442,384A
; CURRENT FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: 09/053,375
; PRIOR FILING DATE: 1998-03-31
; NUMBER OF SEQ ID NOS: 830
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 792
; LENGTH: 1536
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-442-384A-792

Query Match 90.8%; Score 1394; DB 18; Length 1536;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1494; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGGGGGGGGGGGACCACTTGGCTGCTCGTCCCGCGGCACTTGGCTGCTCGT 60
DB 1 GGGGGGGGGGGGACCACTTGGCTGCTCGTCCCGCGGCACTTGGCTGCTCGT 60
QY 61 CCGCGCGCGCACTTGGCTGCTCGTCCCGCGGCACTTGGCTGCTCGT 120
DB 61 CCGCGCGCGCACTTGGCTGCTCGTCCCGCGGCACTTGGCTGCTCGT 120
QY 121 TCGGAGCTGCTCGCGCGCGCTTGGCTGCTCGTCCCGCGGCACTTGGCTGCTCGT 180
DB 121 TCGGAGCTGCTCGCGCGCGCTTGGCTGCTCGTCCCGCGGCACTTGGCTGCTCGT 180
QY 181 CCGCGCGCGCACTTGGCTGCTCGTCCCGCGGCACTTGGCTGCTCGT 240
DB 181 CCGCGCGCGCACTTGGCTGCTCGTCCCGCGGCACTTGGCTGCTCGT 240
QY 241 AGGAAAG 300
DB 241 AGGAAAG 300

QY 301 AGCCTGAGAGATGAATTCCTCTCTGCTGACACCAACGTCGTCTGGAAGAGGTTTG 360
DB 301 AGCCTGAGAGATGAATTCCTCTCTGCTGACACCAACGTCGTCTGGAAGAGGTTTG 360
QY 361 GAGAGATGCTGTGTTTATCAAGAGATCCCAAAATGCTTAAAGAGCTGCTTCAAGGGA 420
DB 361 GAGAGATGCTGTGTTTATCAAGAGATCCCAAAATGCTTAAAGAGCTGCTTCAAGGGA 420
QY 421 GTGAATCTGGGATGCAATGATTCGAGACTTTTGAACAGCTTGGATTTCTCCACC 480
DB 421 GTGAATCTGGGATGCAATGATTCGAGACTTTTGAACAGCTTGGATTTCTCCACC 480
QY 481 AG 540
DB 481 AG 540
QY 541 TACAGAGATGAGATTCAGATTTATTCAGGACAGGAGCTTGAACAACTGCAAGAGAT 600
DB 541 TACAGAGATGAGATTCAGATTTATTCAGGACAGGAGCTTGAACAACTGCAAGAGAT 600
QY 601 GACACCAATCAAAACCAACCTGACAGACAGAGAGATCATGTCGTGATTCAGAG 660
DB 601 GACACCAATCAAAACCAACCTGACAGACAGAGAGATCATGTCGTGATTCAGAG 660
QY 661 GATCTCTCTGATGAGCTGCTGCTTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
DB 661 GATCTCTCTGATGAGCTGCTGCTTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
QY 721 AGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
DB 721 AGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
QY 781 AACATGCGAGCTGAG 840
DB 781 AACATGCGAGCTGAG 840
QY 841 GGTGACTTTTAAACACTTGGAGAGATGCAATATTCCTGATCATGAGAGAGAGAG 900
DB 841 GGTGACTTTTAAACACTTGGAGAGATGCAATATTCCTGATCATGAGAGAGAGAG 900
QY 901 AAAATTCAGCTTCAAG 960
DB 901 AAAATTCAGCTTCAAG 960
QY 961 GAGAAATTTGATGATCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
DB 961 GAGAAATTTGATGATCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
QY 1021 ACTATTAATTAAG 1080
DB 1021 ACTATTAATTAAG 1080
QY 1081 GTCTTTAGGGGTTGGCGTGAATCCGAGGTAAAGTTCTTTTGTCTTAAAGANAAG 1140
DB 1081 GTCTTTAGGGGTTGGCGTGAATCCGAGGTAAAGTTCTTTTGTCTTAAAGANAAG 1140
QY 1141 AACTAGGTCAAAAATCTGTCCGTGACCTTATGATTAATTTTAAAGANTTCCACT 1200
DB 1141 AACTAGGTCAAAAATCTGTCCGTGACCTTATGATTAATTTTAAAGANTTCCACT 1200
QY 1201 GCGAATGTAATCTGTCAGTCTTTCATATAAAGGCTTTGAGTAACTCACTAGAG 1260
DB 1201 GCGAATGTAATCTGTCAGTCTTTCATATAAAGGCTTTGAGTAACTCACTAGAG 1260
QY 1261 GTATCTGACAAATGCTGAGTTATGAACAAAGTGAAGAAAGTAATGTCCTTAG 1320
DB 1261 GTATCTGACAAATGCTGAGTTATGAACAAAGTGAAGAAAGTAATGTCCTTAG 1320
QY 1321 CAATAACATGATGATGCAATTCATCCAGTACTTAAAGAGGTTGATTTTCAC 1380
DB 1321 CAATAACATGATGATGCAATTCATCCAGTACTTAAAGAGGTTGATTTTCAC 1380

Query Match	90.8%;	Score 1394;	DB 1;	Length 1536;
Best Local Similarity	99.9%;	Pred. No. 0;		

Matches 1494; Conservative 0; Mismatches 2; Indels 0; Gaps 0			
Qy	1	GGGGGGGGGGGGGACCACTTGGGCTGCTCCGCTCCGCGGCACTTGGCTGCTCCG	60
Db	1	GGGGGGGGGGGGGACCACTTGGCTGCTCCGCTCCGCGGCACTTGGCTGCTCCG	60
Qy	61	CCCGCCGGGCACTTGGGCTGCTCCGCTCCGCGGCACTTGGCTGCTCCGCG	120
Db	61	CCCGCCGGGCACTTGGGCTGCTCCGCTCCGCGGCACTTGGCTGCTCCGCG	120
Qy	121	TGGGAGCTGGGGCGGGCCCTTGGCCCGCGGCAAGAGGGGAGCGGAGCGCGG	180
Db	121	TGGGAGCTGGGGCGGGCCCTTGGCCCGCGGCAAGAGGGGAGCGGAGCGCGG	180
Qy	181	CCGCGGCAAGGGAGCTGCAATCTGGGGGAGATCCACACATCCGCTGGGGGCTC	240
Db	181	CCGCGGCAAGGGAGCTGCAATCTGGGGGAGATCCACACATCCGCTGGGGGCTC	240
Qy	241	AGGAGAGACGACCCGCAAGGCAACCGGCAACCTGTGCGTATTCGAGATCAGGCGCGTAC	300
Db	241	AGGAGAGACGACCCGCAAGGCAACCGGCAACCTGTGCGTATTCGAGATCAGGCGCGTAC	300
Qy	301	AGCTGAGAGATGATTCCTCTGCTGACCAACAAAGTGTTCTGGAAGGGTGTTTG	360
Db	301	AGCTGAGAGATGATTCCTCTGCTGACCAACAAAGTGTTCTGGAAGGGTGTTTG	360
Qy	361	GAGGAGTTCGTGTGTTTATCAAGGAGTCCAAATGCTTAAAGAGCTGCTTCCAGGGA	420
Db	361	GAGGAGTTCGTGTGTTTATCAAGGAGTCCAAATGCTTAAAGAGCTGCTTCCAGGGA	420
Qy	421	GTGGAATCTGGGATGTCATAGGATCCCGGACCTTTTGGACAGCTGGGATCTCCACG	480
Db	421	GTGGAATCTGGGATGTCATAGGATCCCGGACCTTTTGGACAGCTGGGATCTCCACG	480
Qy	481	AGGAGAGAGGGGACTTGGGCCGATTTAGCTCCAGTGAAGGACATTTTGGGGAGAA	540
Db	481	AGGAGAGAGGGGACTTGGGCCGATTTAGCTCCAGTGAAGGACATTTTGGGGAGAA	540
Qy	541	TACAGAGATAGGATCAGATTATTCAGAGACGGGAGTTGACCACTGCAAGAGTATT	600
Db	541	TACAGAGATAGGATCAGATTATTCAGAGACGGGAGTTGACCACTGCAAGAGTATT	600
Qy	601	GACACCATCAAAACCAACCTTGACGACAGAAATCATATGTCGCTGGATCCAGAA	660
Db	601	GACACCATCAAAACCAACCTTGACGACAGAAATCATATGTCGCTGGATCCAGAA	660
Qy	661	GATCTTCTCTGATGCGGCTGCTCATAGCCATGCTCTGCACTTATATGTTGTAAC	720
Db	661	GATCTTCTCTGATGCGGCTGCTCATAGCCATGCTCTGCACTTATATGTTGTAAC	720
Qy	721	AGTGAAGCTGTCGCGACGATGACAGAGATCGGGAGACATGGGCTCGGTGGCTTTC	780
Db	721	AGTGAAGCTGTCGCGACGATGACAGAGATCGGGAGACATGGGCTCGGTGGCTTTC	780
Qy	781	AACATCGACACTTCGCGCTGCTCAGATGATGATGGCCACATACAGGGCTGAAAGCA	840
Db	781	AACATCGACACTTCGCGCTGCTCAGATGATGATGGCCACATACAGGGCTGAAAGCA	840
Qy	841	GGTGACTTTATACACTTTGGGATGCAATATTTAATCTGAATCATGAGGACCTG	900
Db	841	GGTGACTTTATACACTTTGGGATGCAATATTTAATCTGAATCATGAGGACCTG	900
Qy	901	AAAATTCAGCTTCGAGAGACCCGACCTTTCCAAAGCTCAGGATCTTCCAAAAGTT	960
Db	901	AAAATTCAGCTTCGAGAGACCCGACCTTTCCAAAGCTCAGGATCTTCCAAAAGTT	960
Qy	961	GAGAAAATGATGATCTTCAAAGCTGAGACCTTCAATTTGAAGGGTACATCCGATCCA	1020
Db	961	GAGAAAATGATGATCTTCAAAGCTGAGACCTTCAATTTGAAGGGTACATCCGATCCA	1020
Qy	1021	ACATTTAAATAGGAATGCTGTTTAAAGGCTTTCAAGAGAGCTTGAAGAGATATTTGTA	1080
Db	1021	ACATTTAAATAGGAATGCTGTTTAAAGGCTTTCAAGAGAGCTTGAAGAGATATTTGTA	1080

Db 1021 ACTATTAATAATGGAATGCTGTGTTAGGGTCTTTCAAAAGAGCTNGAAGATATTGTCA 1080
 QY 1081 GTCTTTGGGGGTTGGGCTGATGCCAGGTAAAAGTCTTTTGTCTTAAAGANAAG 1140
 Db 1081 GTCTTTGGGGGTTGGGCTGATGCCAGGTAAAAGTCTTTTGTCTTAAAGANAAG 1140
 QY 1141 AACTAGGTCAAAAATCTGTCCGTGACCTATCAGTTATTAATTTTAAAGATGTTGCCACT 1200
 Db 1141 AACTAGGTCAAAAATCTGTCCGTGACCTATCAGTTATTAATTTTAAAGATGTTGCCACT 1200
 QY 1201 GGCAAAATGTAAGTGTGCCAGTTCTTTCCATTAATAAAGCTTTGAGTTAACTCACTGAG 1260
 Db 1201 GGCAAAATGTAAGTGTGCCAGTTCTTTCCATTAATAAAGCTTTGAGTTAACTCACTGAG 1260
 QY 1261 GTATCTGACAACTGAGTTATTAAGCAAAAGTGAAGAAATGAATGATGCTCTAG 1320
 Db 1261 GTATCTGACAACTGAGTTATTAAGCAAAAGTGAAGAAATGAATGATGCTCTAG 1320
 QY 1321 CAAAACATGATGATGATTCATTCATCCACGTAATTATAAGAGGTTGTGAATTTCA 1380
 Db 1321 CAAAACATGATGATGATTCATTCATCCACGTAATTATAAGAGGTTGTGAATTTCA 1380
 QY 1381 AAGCTATTTTGGATATTTTGAATTTTGAATTTTGAATTTTCAAGCTATTCCTCAAT 1440
 Db 1381 AAGCTATTTTGGATATTTTGAATTTTGAATTTTGAATTTTCAAGCTATTCCTCAAT 1440
 QY 1441 CTGAGGAGCTGATGAACACATCATCATGATGATGATGATGATGATGATGATGATGAT 1500
 Db 1441 CTGAGGAGCTGATGAACACATCATCATGATGATGATGATGATGATGATGATGATGAT 1500
 QY 1501 TGTTTTATGTTGCTTATTAATAAGATGTTCTGC 1536
 Db 1501 TGTTTTATGTTGCTTATTAATAAGATGTTCTGC 1536

RESULT 5

PCT-US02-18947-556

Sequence 556, Application PC/TUS0218947

GENERAL INFORMATION:

APPLICANT: Rosetta Inpharmatics

TITLE OR INVENTION: Diagnosis and Prognosis of Breast Cancer Patients

FILE REFERENCE: 9301-175-228

CURRENT APPLICATION NUMBER: PCT/US02/18947

CURRENT FILING DATE: 2002-06-14

PRIOR APPLICATION NUMBER: 60/380,770

PRIOR FILING DATE: 2002-05-14

NUMBER OF SEQ ID NOS: 2699

SEQ ID NO 556

LENGTH: 1536

TYPE: DNA

ORGANISM: Homo sapiens

PUBLICATION INFORMATION:

DATABASE ACCESSION NUMBER: NM_001071

DATABASE ENTRY DATE: 2001-06-18

PCT-US02-18947-556

Query Match 90.8%; Score 1394; DB 1; Length 1536;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1494; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGGGGGGGGGGGACCTTGGCTGGCTCCGCTCCGCGCGACCACTTGGCTGGCTCCGCT 60
 Db 1 GGGGGGGGGGGGACCTTGGCTGGCTCCGCTCCGCGCGACCACTTGGCTGGCTCCGCT 60
 QY 61 CCCCCCGCCGCACTTGGCTGGCTCCGCTCCGCGCGCGCGCGCGCACTTGGCTGGCTCCGCT 120
 Db 61 CCCCCCGCCGCACTTGGCTGGCTCCGCTCCGCGCGCGCGCGCGCGCACTTGGCTGGCTCCGCT 120
 QY 121 TCGAGAGCTGCGCGCGCGCTTGGCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 180
 Db 121 TCGAGAGCTGCGCGCGCGCTTGGCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 180
 QY 181 CCGCGCGCGCGCGCGCTGCGAGTCTGGGGGCGAGATCCACAACATCTCTCCGCTGCGCGCTC 240

Db 181 CCGCGCGCGCGCGCGCTGCGAGTCTGGGGGCGAGTCGCAACATCTCCGCTCGCGCGCTC 240
 QY 241 AGGAGAGAGAGCGAGCTAC 300
 Db 241 AGGAGAGAGAGCGAGCTAC 300
 QY 301 AGCTGAGAGAGAGATTCCT 360
 Db 301 AGCTGAGAGAGAGATTCCT 360
 QY 361 GAGAGAGAGAGAGATTCCT 420
 Db 361 GAGAGAGAGAGAGATTCCT 420
 QY 421 GTGAATCTGGAGATGCCAATGATGCCAATGATGCCAATGATGCCAATGATGCCAATGATGCC 480
 Db 421 GTGAATCTGGAGATGCCAATGATGCCAATGATGCCAATGATGCCAATGATGCCAATGATGCC 480
 QY 481 AGAGAGAGAGAGAGCTTGGGCGCGAGTTATGCTTCAAGTGAAGCATTTTGGGCGAGAA 540
 Db 481 AGAGAGAGAGAGAGCTTGGGCGCGAGTTATGCTTCAAGTGAAGCATTTTGGGCGAGAA 540
 QY 541 TACAGAGATGAGATTCAGATTATTCAGAGAGAGAGAGAGAGATTCAGAGAGAGAGATTC 600
 Db 541 TACAGAGATGAGATTCAGATTATTCAGAGAGAGAGAGAGAGATTCAGAGAGAGAGATTC 600
 QY 601 GACACCATCAAAACCAACCTCGAGAGAGAGAGAGAGAGATTCAGAGAGAGAGATTCAGAG 660
 Db 601 GACACCATCAAAACCAACCTCGAGAGAGAGAGAGAGAGATTCAGAGAGAGAGATTCAGAG 660
 QY 661 GATCTTCTCTGATGAGCGCTGCTCCATGCGATGCCCTGCGCACTTCTATGTTGATG 720
 Db 661 GATCTTCTCTGATGAGCGCTGCTCCATGCGATGCCCTGCGCACTTCTATGTTGATG 720
 QY 721 AGTGAAGCTGTCTGCGAGCTGTACAGAGATGCGAGAGAGAGAGAGAGATGCGAGCTTTTC 780
 Db 721 AGTGAAGCTGTCTGCGAGCTGTACAGAGATGCGAGAGAGAGAGAGAGATGCGAGCTTTTC 780
 QY 781 AACATGCGAGCTAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
 Db 781 AACATGCGAGCTAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
 QY 841 GGTGACCTTATACACTTTGGGAGAGAGAGAGAGAGAGATTCAGAGAGAGAGAGAGAGAG 900
 Db 841 GGTGACCTTATACACTTTGGGAGAGAGAGAGAGAGAGATTCAGAGAGAGAGAGAGAGAG 900
 QY 901 AAATTCAGCTTGAAG 960
 Db 901 AAATTCAGCTTGAAG 960
 QY 961 GAGAAATGATGACTTCAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
 Db 961 GAGAAATGATGACTTCAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
 QY 1021 ACTATTAATAATGGAATGCTGTGTTAGGGTCTTTCAAAAGAGCTNGAAGATATTGTCA 1080
 Db 1021 ACTATTAATAATGGAATGCTGTGTTAGGGTCTTTCAAAAGAGCTNGAAGATATTGTCA 1080
 QY 1081 GTCTTTGGGGGTTGGGCTGATGCCAGGTAAAAGTCTTTTGTCTTAAAGANAAG 1140
 Db 1081 GTCTTTGGGGGTTGGGCTGATGCCAGGTAAAAGTCTTTTGTCTTAAAGANAAG 1140
 QY 1141 AACTAGGTCAAAAATCTGTCCGTGACCTATCAGTTATTAATTTTAAAGATGTTGCCACT 1200
 Db 1141 AACTAGGTCAAAAATCTGTCCGTGACCTATCAGTTATTAATTTTAAAGATGTTGCCACT 1200
 QY 1201 GGCAAAATGTAAGTGTGCCAGTTCTTTCCATTAATAAAGCTTTGAGTTAACTCACTGAG 1260
 Db 1201 GGCAAAATGTAAGTGTGCCAGTTCTTTCCATTAATAAAGCTTTGAGTTAACTCACTGAG 1260
 QY 1261 GTATCTGACAACTGAGTTATTAAGCAAAAGTGAAGAAATGAATGATGCTCTAG 1320

Qy 1081 GCTTTAGGGGTTGGCTGATGCGAGTAAGTCTTTTCTCTTAAAGAAAG 1140
Db 1081 GCTTTAGGGGTTGGCTGATGCGAGTAAGTCTTTTCTCTTAAAGAAAG 1140
Qy 1141 AACTAGGTCAAAAATCTGTCCGTGACCTATCAGTTATTATTAAAGATGTGCCACT 1200
Db 1141 AACTAGGTCAAAAATCTGTCCGTGACCTATCAGTTATTATTAAAGATGTGCCACT 1200
Qy 1201 GGGCAATGTATCTGTGCGAGTCTTTCTCATATAAAGCTTTAGATTACTGACGAG 1260
Db 1201 GGGCAATGTATCTGTGCGAGTCTTTCTCATATAAAGCTTTAGATTACTGACGAG 1260
Qy 1261 GTATCTGACAAATGTGAGGTTATGACAAAGTGAAGAAATGAATGAGTCTTAC 1320
Db 1261 GTATCTGACAAATGTGAGGTTATGACAAAGTGAAGAAATGAATGAGTCTTAC 1320
Qy 1321 CAAAACATGATGTGATTCATCCACGACTTAAAGAGTGTGTAATTCAC 1380
Db 1321 CAAAACATGATGTGATTCATCCACGACTTAAAGAGTGTGTAATTCAC 1380
Qy 1381 AAGCTATTTTGAATATTTTGAATATTTTGAATTTTCAAGCTATTCCTCAAT 1440
Db 1381 AAGCTATTTTGAATATTTTGAATATTTTGAATTTTCAAGCTATTCCTCAAT 1440
Qy 1441 CTGAGGAGCTGATGACACCATGATCATGATGAGTGTGTTAATCTTAAAGT 1500
Db 1441 CTGAGGAGCTGATGACACCATGATCATGATGAGTGTGTTAATCTTAAAGT 1500
Qy 1501 TGTATTATGTGTCTAATAAAGAGTCTGAC 1536
Db 1501 TGTATTATGTGTCTAATAAAGAGTCTGAC 1536

RESULT 4
US-09-963-333-6
; Sequence 6: Application US/09963333
; GENERAL INFORMATION:
; APPLICANT: Stanton, Jr., Vincent P.
; TITLE OF INVENTION: THYMIDINE SYNTHASE GENE SEQUENCE VARIANCES
; TITLE OF INVENTION: HAVING UTILITY IN DETERMINING THE TREATMENT
; TITLE OF INVENTION: OF DISEASE
; FILE REFERENCE: 11926-015002
; CURRENT APPLICATION NUMBER: US/09/963,333
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/658,659
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 09/596,033
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 09/357,743
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 09/357,024
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: 60/093,484
; PRIOR FILING DATE: 1998-07-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1536
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1066
; OTHER INFORMATION: n = t or c
; NAME/KEY: misc_feature
; LOCATION: 1136
; OTHER INFORMATION: n = a or g
; NAME/KEY: misc_feature
; LOCATION: 1497
; OTHER INFORMATION: n = t or a
US-09-963-333-6
Query Match 100.0%; Score 1536; DB 36; length 1536;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 1536; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GGGGGGGGGGGGCACTTGGCTGCTGCTCCGCGCGGCACTTGGCTGCTCGT 60
Db 1 GGGGGGGGGGGGCACTTGGCTGCTGCTCCGCGCGGCACTTGGCTGCTCGT 60
Qy 61 CCCGGCGGCACTTGGCTGCTGCTCCGCGCGGCACTTGGCTGCTCGT 120
Db 61 CCCGGCGGCACTTGGCTGCTGCTCCGCGCGGCACTTGGCTGCTCGT 120
Qy 121 TCGAGCTGCGCGCGGCTGCTGCTCCGCGCGGCACTTGGCTGCTCGT 180
Db 121 TCGAGCTGCGCGCGGCTGCTGCTCCGCGCGGCACTTGGCTGCTCGT 180
Qy 181 CCGCGCGGCACTTGGCTGCTGCTCCGCGCGGCACTTGGCTGCTCGT 240
Db 181 CCGCGCGGCACTTGGCTGCTGCTCCGCGCGGCACTTGGCTGCTCGT 240
Qy 241 AGAAGAGCAAGCGGCACTTGGCTGCTGCTCCGCGCGGCACTTGGCTGCTCGT 300
Db 241 AGAAGAGCAAGCGGCACTTGGCTGCTGCTCCGCGCGGCACTTGGCTGCTCGT 300
Qy 301 AAGCTGAGATGAAATCCCTGCTGCTGCAACCAAGCTGCTGGAAGGCTTTTG 360
Db 301 AAGCTGAGATGAAATCCCTGCTGCTGCAACCAAGCTGCTGGAAGGCTTTTG 360
Qy 361 GAGAGGTTGCTGCTGCTTATCAAGGATCAAGATCTTAAAGCTGCTTCAAGGGA 420
Db 361 GAGAGGTTGCTGCTGCTTATCAAGGATCAAGATCTTAAAGCTGCTTCAAGGGA 420
Qy 421 GTGAATCTGGATGCAATGATCCGAGACTTTTGAAGCTGCTGCTTCAAGC 480
Db 421 GTGAATCTGGATGCAATGATCCGAGACTTTTGAAGCTGCTGCTTCAAGC 480
Qy 481 AGAAGAGGAGGAGCTTGGGCGGCACTTATGCTTCAAGGAGCTTTGGGAGGA 540
Db 481 AGAAGAGGAGGAGCTTGGGCGGCACTTATGCTTCAAGGAGCTTTGGGAGGA 540
Qy 541 TACGAGATGAGATGATATGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600
Db 541 TACGAGATGAGATGATATGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600
Qy 601 GACACATCAAAAACAACCTGACAGAGAGATCATATGCTGCTGAGATCCAGA 660
Db 601 GACACATCAAAAACAACCTGACAGAGAGATCATATGCTGCTGAGATCCAGA 660
Qy 661 GATCTTCTGATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Db 661 GATCTTCTGATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Qy 721 AGTGAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db 721 AGTGAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Qy 781 AACATCGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Db 781 AACATCGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Qy 841 GGTGACTTATACACTTTGGAGATGACATATTTTAACTGATCAATGAGCCACTG 900
Db 841 GGTGACTTATACACTTTGGAGATGACATATTTTAACTGATCAATGAGCCACTG 900
Qy 901 AAAATTGAGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
Db 901 AAAATTGAGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
Qy 961 GAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
Db 961 GAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
Qy 1021 ACTATTAAATGGAATGAGCTGTTAGGCTGCTTCAAGAGGCTGAGAGATATTTGCA 1080

Matches 1536; Conservative 0; Mismatches

QY	1	GGGGGGGGGGGGGACCACTTGAGCTGCGCTCCCGGCGCGGACCTTGGCTGCTCGCT	60
Db	1	GGGGGGGGGGGGGACCACTTGAGCTGCTCGCTCCCGGCGCGGACCTTGGCTGCTCGCT	60
QY	61	CGCGCGCGGCGCACTTGCGCTGCGCTCGCTCCCGGCGCGGACCTTGGCTGCTCGCT	120
Db	61	CGCGCGCGGCGCACTTGCGCTGCGCTCGCTCCCGGCGCGGACCTTGGCTGCTCGCT	120
QY	121	TGGAGCTGCGCGCGCGCTTGCGCTCCCGGCGCGGACCTTGGCTGCTCGCT	180
Db	121	TGGAGCTGCGCGCGCGCTTGCGCTCCCGGCGCGGACCTTGGCTGCTCGCT	180
QY	181	CCCGCGCGCGGAGCTGCAAGTACTTGCGGAGATCCACACATCTCGCTGCGCGCT	240
Db	181	CCCGCGCGCGGAGCTGCAAGTACTTGCGGAGATCCACACATCTCGCTGCGCGCT	240
QY	241	AGGAGGAGCGACCGCGCGCGGCGACCGCGACCTTGCTGCTGATTCGGGCACTGAGCGCGCTAC	300
Db	241	AGGAGGAGCGACCGCGCGCGGCGACCGCGACCGCTTGCTGCTGATTCGGGCACTGAGCGCGCTAC	300
QY	301	AGCTTGAGAGTGAATTCCTCTGCTGCAACCAAGTGTGTCGGAAGGGTGTTTG	360
Db	301	AGCTTGAGAGTGAATTCCTCTGCTGCAACCAAGTGTGTCGGAAGGGTGTTTG	360
QY	361	GAGGAGTGTGCTGTGATTCCTTGAGGAGTCCAGCAATCTGAAGAGCTGTCTCCAGAGGA	420
Db	361	GAGGAGTGTGCTGTGATTCCTTGAGGAGTCCAGCAATCTGAAGAGCTGTCTCCAGAGGA	420
QY	421	GTCGAATCTGGGAGTGCATATGATCCGAGACTTTTGAGAGCTTGAGATTCTGCACC	480
Db	421	GTCGAATCTGGGAGTGCATATGATCCGAGACTTTTGAGAGCTTGAGATTCTGCACC	480
QY	481	AGGAGAGAGGGGACTTGCGGCGCGATTATGCGTTTCCAGTGAGGCGATTTGGGGCGAA	540
Db	481	AGGAGAGAGGGGACTTGCGGCGCGATTATGCGTTTCCAGTGAGGCGATTTGGGGCGAA	540
QY	541	TACAGAGATATGGAATCGATATTTGAGGACAGGAGTTGACCACTGGAAGAGTATTT	600
Db	541	TACAGAGATATGGAATCGATATTTGAGGACAGGAGTTGACCACTGGAAGAGTATTT	600
QY	601	GACACATCAAAAACCACTCGACGACAGAGAAATCATGATGCGCTGGAAATCCAGA	660
Db	601	GACACATCAAAAACCACTCGACGACAGAGAAATCATGATGCGCTGGAAATCCAGA	660
QY	661	GATTCCTCTGATGCGCGCTCGATCGACATGCGCTCTGCGAGTTCTATATGTGTGAC	720
Db	661	GATTCCTCTGATGCGCGCTCGATCGACATGCGCTCTGCGAGTTCTATATGTGTGAC	720
QY	721	AGTAGAGTGTCTGCGCACTGTACAGAGATGGGAGACATGGGCTCGGTTGCTTTTC	780
Db	721	AGTAGAGTGTCTGCGCACTGTACAGAGATGGGAGACATGGGCTCGGTTGCTTTTC	780
QY	781	AACATCGCAGCTACGCGCTGTCTCAGTACATGATTTGGGCACTCAGGAGCTGAAAGCA	840
Db	781	AACATCGCAGCTACGCGCTGTCTCAGTACATGATTTGGGCACTCAGGAGCTGAAAGCA	840
QY	841	GGTGACTTTATACACACTTTGGAGATGCAATATTTACTGGAATCACAATCGAGCTACTG	900
Db	841	GGTGACTTTATACACACTTTGGAGATGCAATATTTACTGGAATCACAATCGAGCTACTG	900
QY	901	AAATTCAGCTTCACGCGAGACCGACACTTTCCAAAGCTCAGAGATCTTGAGAAAGTT	960
Db	901	AAATTCAGCTTCACGCGAGACCGACACTTTCCAAAGCTCAGAGATCTTGAGAAAGTT	960
QY	961	GAGAAATTTGATGACTTCAAGCTGGAAGCTTTGAGTGAAGGATTCATCCGATCCA	1020
Db	961	GAGAAATTTGATGACTTCAAGCTGGAAGCTTTGAGTGAAGGATTCATCCGATCCA	1020
QY	1021	ACTATTAATTTGAAATGCGCTTGAGGGTCTTCAAGAGCTGAGAGATTTGTCCA	1080
Db	1021	ACTATTAATTTGAAATGCGCTTGAGGGTCTTCAAGAGCTGAGAGATTTGTCCA	1080

GGGGGACCAAGCCAGCCCGCCGACCAAGCCAGCCAGT 60
 1
 22

Db	1	GGGGGGGGGGGGACCACTTGGCTTCCTCCGTCCGGCCGACCTTGGACTTCCTCCGT	60
Oy	61	CCCGCGCGCCACTTTCGCTTCCTCCGTCCCGCCCGCGCGCCATGCTGTGGCCGGC	120
Db	61	CCCGCGCGCCACTTGGCTTCCTCCGTCCCGCCCGCGCGCCATGCTGTGGCCGGC	120
Oy	121	TGGAGCTGCCGCGCGGACTTGGCCCGCCCGCGCCAGAGCGCGACCGCGAGCGCT	180
Db	121	TGGAGCTGCCGCGCGGACTTGGCCCGCCCGCGCCAGAGCGCGACCGCGAGCGCT	180
Oy	181	CCGCTGCACCGGGAGCTGCAGTACTTGGGGCAGATCCACATCTTCGCTCGCGCGTC	240
Db	181	CCGCGCGACGGGGAGCTGCAGTACTTGGGGCAGATCCACATCTTCGCTCGCGCGTC	240
Oy	241	AGGAGAGACGACCGGCAAGCGGACCCCTGTGGGATTCGGCATGAGGCGCGCTAC	300
Db	241	AGGAGAGACGACCGGCAAGCGGACCCCTGTGGGATTCGGCATGAGGCGCGCTAC	300
Oy	301	AGCTGAGAGATGATTCCTCTGTGCAGCAACAAACGTGTCTGGAGGGGTGTTG	360
Db	301	AGCTGAGAGATGATTCCTCTGTGCAGCAACAAACGTGTCTGGAGGGGTGTTG	360
Oy	361	GAGAGTGTGCTGTCTTATCAAGGATCCACAAATGCTAAAGCTGTCCAAAGGA	420
Db	361	GAGAGTGTGCTGTCTTATCAAGGATCCACAAATGCTAAAGCTGTCTTCAAGGA	420
Oy	421	GTCGAAATCTGGGATGCGAATGATCCCGAGCTTTTGGACAGCTGGGATTCCTCAC	480
Db	421	GTCGAAATCTGGGATGCGAATGATCCCGAGCTTTTGGACAGCTGGGATTCCTCAC	480
Oy	481	AGGAGAAAGGGGACTTGGGCCAGTTATGCTTCAGATGAGGCAATTTGGGCGAGA	540
Db	481	AGGAGAAAGGGGACTTGGGCCAGTTATGCTTCAGATGAGGCAATTTGGGCGAGA	540
Oy	541	TACAGAGAATAGAACACAAATTAATCAAGACAGGGAGTTGACCAATCGAAAGGTAT	600
Db	541	TACAGAGAATAGAACACAAATTAATCAAGACAGGGAGTTGACCAATCGAAAGGTAT	600
Oy	601	GACACATCAAAAACCAACCTCGACGACAGAGAAATCATATGTCGACTTGGATCCAA	660
Db	601	GACACATCAAAAACCAACCTCGACGACAGAGAAATCATATGTCGACTTGGATCCAA	660
Oy	661	GATCTTCCTGTGATGGCGCTGACCTTCATGCGCATGCGCTTCGCAATTCATGTGTG	720
Db	661	GATCTTCCTGTGATGGCGCTGACCTTCATGCGCATGCGCTTCGCAATTCATGTGTG	720
Oy	721	AGTGAAGCTGTCGCGCAGTGTACAGAGATGGGAGCATGGGCTGTGGTGTCTTC	780
Db	721	AGTGAAGCTGTCGCGCAGTGTACAGAGATGGGAGCATGGGCTGTGGTGTCTTC	780
Oy	781	AACATGCGAGCTAGGCGCTGCTCAAGTACAGATGTGGACATCCAGCGGCTGAAAG	840
Db	781	AACATGCGAGCTAGGCGCTGCTCAAGTACAGATGTGGACATCCAGCGGCTGAAAG	840
Oy	841	GGTGACTTTATACACTTTGGGAGTGCACATTTACTGTAATCAATGAGCCACTG	900
Db	841	GGTGACTTTATACACTTTGGGAGTGCACATTTACTGTAATCAATGAGCCACTG	900
Oy	901	AAATTCAGCTTCAGCGGAACCCGACCTTCCCAAGCTCGAGATTCCTGAAAAGTT	960
Db	901	AAATTCAGCTTCAGCGGAACCCGACCTTCCCAAGCTCGAGATTCCTGAAAAGTT	960
Oy	961	GAGGAAATGATGACTTCAAAGCTGAAAGCTTCAGATGAAAGGTAAATCCGATCCA	1020
Db	961	GAGGAAATGATGACTTCAAAGCTGAAAGCTTCAGATGAAAGGTAAATCCGATCCA	1020
Oy	1021	ACTATTTAAATGGAATGCGCTTTGAGGTGCTTCAAGGAGCTGGAAGGATTTGTCA	1080
Db	1021	ACTATTTAAATGGAATGCGCTTTGAGGTGCTTCAAGGAGCTGGAAGGATTTGTCA	1080
Oy	1081	GTCCTTAGGGGTGGGCTGAGTCCGAGGTAAAGTCTTTTGCCTTAAAGGANAAG	1140

22 1382 90 0 1829 29 US-09-726-806-5390 Sequence 5390, Ap
 23 1382 90 0 1829 29 US-09-726-807-3895 Sequence 3895, Ap
 24 1346 87 6 3298 63 US-60-195-106-200 Sequence 200, App
 25 1346 87 6 3298 63 US-60-217-674-28 Sequence 28, App1
 26 1346 87 6 3298 71 US-60-278-258-674 Sequence 674, App
 27 1280 83 3 1633 61 US-60-172-373-69 Sequence 69, App1
 28 1236 80 5 1721 30 US-09-760-475-391 Sequence 391, App
 29 1230 80 5 1721 30 US-09-760-475-391 Sequence 391, App
 30 942 61 3 942 18 US-10-084-817-332 Sequence 332, App
 31 942 61 3 942 24 US-09-457-205-29 Sequence 29, App1
 32 942 61 3 942 24 US-09-631-275-143 Sequence 143, App
 33 938 61 1 1540 22 PCT-US01-14827-3952 Sequence 3952, Ap
 34 938 61 1 1540 22 US-09-577-408-4887 Sequence 4887, Ap
 35 882 57 4 1395 30 US-09-760-475-1516 Sequence 1516, Ap
 36 882 57 4 1395 30 US-09-760-485-46 Sequence 46, App1
 37 882 57 4 1395 42 US-10-216-436-46 Sequence 46, App1
 38 537 35 0 18597 25 US-09-658-659-8 Sequence 8, App1
 39 537 35 0 18597 36 US-09-962-677-8 Sequence 8, App1
 40 537 35 0 18597 36 US-09-963-333-8 Sequence 8, App1
 41 536 34 9 1539 3 PCT-US01-08631-20764 Sequence 20764, A
 42 532 34 6 532 33 US-09-878-178-2148 Sequence 2148, Ap
 43 532 34 6 532 38 US-10-046-935-2148 Sequence 2148, Ap
 44 532 34 6 532 40 US-10-146-502-2148 Sequence 2148, Ap
 45 487 31 7 18596 22 US-09-577-266-11 Sequence 11, App1

ALIGNMENTS

RESULT 1
 US-09-658-659-6

; Sequence 6, Application US/09658659

; GENERAL INFORMATION:
 ; APPLICANT: Stanton, Jr., Vincent P.
 ; TITLE OF INVENTION: GENE SEQUENCE VARIANCE IN GENES RELATED
 ; TO POLAR METABOLISM HAVING UTILITY IN DETERMINING THE
 ; TITLE OF INVENTION: TREATMENT OF DISEASE
 ; FILE REFERENCE: 11926-015001
 ; CURRENT APPLICATION NUMBER: US/09/658,659
 ; CURRENT FILING DATE: 2000-09-08
 ; PRIOR APPLICATION NUMBER: 09/596,033
 ; PRIOR FILING DATE: 2000-06-15
 ; PRIOR APPLICATION NUMBER: 09/357,743
 ; PRIOR FILING DATE: 1999-07-20
 ; PRIOR APPLICATION NUMBER: 09/357,024
 ; PRIOR FILING DATE: 1999-07-19
 ; PRIOR APPLICATION NUMBER: 60/093,484
 ; PRIOR FILING DATE: 1998-07-20
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 6
 ; LENGTH: 1536
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: 1066
 ; OTHER INFORMATION: n = t or c
 ; NAME/KEY: misc_feature
 ; LOCATION: 1136
 ; OTHER INFORMATION: n = a or g
 ; NAME/KEY: misc_feature
 ; LOCATION: 1497
 ; OTHER INFORMATION: n = t or a
 ; US-09-658-659-6

Query Match 100.0%; Score 1536; DB 25; Length 1536;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1536; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGACACTTGGCTGCTCCGTCGCGCGGACCACTTGGCTGCTCCGCT
 Db 1 GGGGGGGGGGGGACACTTGGCTGCTCCGTCGCGCGGACCACTTGGCTGCTCCGCT 60

Qy 61 CCCGCGCGGACCTTGGCTGCTCCGTCGCGCGGACCACTTGGCTGCTCCGCT 120
 Db 61 CCCGCGCGGACCTTGGCTGCTCCGTCGCGCGGACCACTTGGCTGCTCCGCT 120
 Qy 121 TCGAGAGTTCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
 Db 121 TCGAGAGTTCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
 Qy 181 CCGCGCGGAGGAGTTCGAGTACCTTGGGAGATCCAGATCTCGGTCGCGCTC 240
 Db 181 CCGCGCGGAGGAGTTCGAGTACCTTGGGAGATCCAGATCTCGGTCGCGCTC 240
 Qy 241 AGGAGAGGAGCGGACCGGCGGACCGGCGGACCGGCGGATTCGAGTTCGAGT 300
 Db 241 AGGAGAGGAGCGGACCGGCGGACCGGCGGACCGGCGGATTCGAGTTCGAGT 300
 Qy 301 AGCTTGAGAGTGAATTCCTCTGCTGACCAACAGTGTCTGGAAGGCTTTTG 360
 Db 301 AGCTTGAGAGTGAATTCCTCTGCTGACCAACAGTGTCTGGAAGGCTTTTG 360
 Qy 361 GAGAGGTTGCTGCTGCTTATTCAGAGGATCCAAATGCTTAAAGCTCTTCAAGGA 420
 Db 361 GAGAGGTTGCTGCTGCTTATTCAGAGGATCCAAATGCTTAAAGCTCTTCAAGGA 420
 Qy 421 GTGAATTCGAGATCCCAATGATCCGAGATCTTTTGGAGAGCTGGGATTCAC 480
 Db 421 GTGAATTCGAGATCCCAATGATCCGAGATCTTTTGGAGAGCTGGGATTCAC 480
 Qy 481 AGAGAGAGGAGGAGCTTGGGCGGAGTTTATGAGGAGGAGGAGGAGGAGGAG 540
 Db 481 AGAGAGAGGAGGAGCTTGGGCGGAGTTTATGAGGAGGAGGAGGAGGAGGAG 540
 Qy 541 TACGAGATGAGATCAGATTAATCAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600
 Db 541 TACGAGATGAGATCAGATTAATCAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600
 Qy 601 GACACATCAAAACCAACCTGACGACGAGAGAGATCATGTCGCTTGAATCAGA 660
 Db 601 GACACATCAAAACCAACCTGACGACGAGAGAGATCATGTCGCTTGAATCAGA 660
 Qy 661 GATCTTCCTGATGAGGAGGCTGCTCAGTCCATGCTTGGCACTTATGAGTGAAC 720
 Db 661 GATCTTCCTGATGAGGAGGCTGCTCAGTCCATGCTTGGCACTTATGAGTGAAC 720
 Qy 721 AGTGAAGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
 Db 721 AGTGAAGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
 Qy 781 AACATGCGAGCTTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
 Db 781 AACATGCGAGCTTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
 Qy 841 GGTGACTTATACCACTTGGGAGTGAATTAATTAATTAATTAATTAATTAATTA 900
 Db 841 GGTGACTTATACCACTTGGGAGTGAATTAATTAATTAATTAATTAATTAATTA 900
 Qy 901 AAAATTCAGCTTACGAGAGACCAAGCTTCCCAAGCTGAGATTCGAAAGTT 960
 Db 901 AAAATTCAGCTTACGAGAGACCAAGCTTCCCAAGCTGAGATTCGAAAGTT 960
 Qy 961 GAGAAATTTGATGACTTCAAGCTGAGATTCAGATTTGAAGGATCAATCCGATCA 1020
 Db 961 GAGAAATTTGATGACTTCAAGCTGAGATTCAGATTTGAAGGATCAATCCGATCA 1020
 Qy 1021 ACTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1080
 Db 1021 ACTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1080
 Qy 1081 GTCTTTAGGAGGTTGGCTGAGTCCGAGGATTAAGTTCTTTTGTCTTAAAGAAAG 1140
 Db 1081 GTCTTTAGGAGGTTGGCTGAGTCCGAGGATTAAGTTCTTTTGTCTTAAAGAAAG 1140

QY 961 GAGAAATGATGACTTCAAAAGCTGAAAGCTTCAGATTGAAGGATCAATCCGACATCA 1020
 DB 1567 GAGAAATGATGACTTCAAAAGCTGAAAGCTTCAGATTGAAGGATCAATCCGACATCA 1508
 QY 1021 ACTATTAAATGAAATGCTGTTTAGGGTCTTTCAAAGAGCTGAAAGATATTGTCA 1080
 DB 1507 ACTATTAAATGAAATGCTGTTTAGGGTCTTTCAAAGAGCTGAAAGATATTGTCA 1448
 QY 1081 GTCTTTAGGGGTTTGGCTGATGCGAGGAAAGTCTTTTGTCT 1126
 DB 1447 GTCTTTAGGGGTTTGGCTGATGCGAGGAAAGTCTTTTGTCT 1402

RESULT 14

US-09-724-676A-34548/c
 ; Sequence 34548, Application US/09724676A
 ; GENERAL INFORMATION:
 ; APPLICANT: Compugen LTD
 ; TITLE OF INVENTION: Variants of alternative splicing
 ; FILE REFERENCE: 129181.4 Compugen
 ; CURRENT APPLICATION NUMBER: US/09/724,676A
 ; CURRENT FILING DATE: 2000-11-28
 ; NUMBER OF SEQ ID NOS: 97222
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 34548
 ; LENGTH: 2527
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (41)..(41)
 ; OTHER INFORMATION: n is a,c,g, or t
 US-09-724-676A-34548

Query Match 66.7%; Score 1024; DB 5; Length 2527;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1124; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGGGGGGGGGGGAGACCACTTGGCTGCTCCGCGCGCGCACTTGGCTGCTCCGT 60
 DB 2527 GGGGGGGGGGGGAGACCACTTGGCTGCTCCGCGCGCGCACTTGGCTGCTCCGT 2468
 QY 61 CCGCGCGCGCACTTGGCTGCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
 DB 2467 CCGCGCGCGCACTTGGCTGCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2408
 QY 121 TCGAGCTGCG 180
 DB 2407 TCGAGCTGCG 2348
 QY 181 CCGCGCGCGAGGAGCTGCACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240
 DB 2347 CCGCGCGCGAGGAGCTGCACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2288
 QY 241 AGGAAGAGACCCGACCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
 DB 2287 AGGAAGAGACCCGACCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2228
 QY 301 AGCTTGAAGATGAATTCCTCTGCTGACCAACAAAGTGTCTTGAAGGGGTGTTTG 360
 DB 2227 AGCTTGAAGATGAATTCCTCTGCTGACCAACAAAGTGTCTTGAAGGGGTGTTTG 2168
 QY 361 GAGGAGTTCCTGCTGTTTATCAAGGATCCAAATGCTAAAGAGCTCTTCCAAAGGGA 420
 DB 2167 GAGGAGTTCCTGCTGTTTATCAAGGATCCAAATGCTAAAGAGCTCTTCCAAAGGGA 2108
 QY 421 GTGAAATCTGGGATGCCAATGATCCCGAAGTTTGTGAACAGCTGGGATCTTCCACC 480
 DB 2107 GTGAAATCTGGGATGCCAATGATCCCGAAGTTTGTGAACAGCTGGGATCTTCCACC 2048
 QY 481 AGAAGAGAGGAGCTTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 540
 DB 2047 AGAAGAGAGGAGCTTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1988

QY 541 TACAGATATGAAATCAGATTATTTCAGAGCAGGAGTGTGACCAACTGCAAAAGATGATT 600
 DB 1987 TACAGATATGAAATCAGATTATTTCAGAGCAGGAGTGTGACCAACTGCAAAAGATGATT 1928
 QY 601 GAACCATCAAAACCACTTGACGACAGAGATTCATGATGTCGCTTGAATCCAGAA 660
 DB 1927 GAACCATCAAAACCACTTGACGACAGAGATTCATGATGTCGCTTGAATCCAGAA 1868
 QY 661 GATCTTCCTCTGATGAGCGCTGCTCCATGCGCATGCCCTCTGCGCATGTTCTATGTTGTAAC 720
 DB 1867 GATCTTCCTCTGATGAGCGCTGCTCCATGCGCATGCCCTCTGCGCATGTTCTATGTTGTAAC 1808
 QY 721 AGTAGCTGTCTCTGACGCTGTATCCAGAGATGCGGAGCATGCGGCTCGGTGCTTTC 780
 DB 1807 AGTAGCTGTCTCTGACGCTGTATCCAGAGATGCGGAGCATGCGGCTCGGTGCTTTC 1748
 QY 781 AACATGCGAGGATGAGCGCTGCTCAAGTATGATGTTGGGACATCAAGCGGCTGAGACCA 840
 DB 1747 AACATGCGAGGATGAGCGCTGCTCAAGTATGATGTTGGGACATCAAGCGGCTGAGACCA 1688
 QY 841 GGTGACTTAAACACACTTGGGAGATGACATATTACCTGATCAATCAAGCCACTG 900
 DB 1687 GGTGACTTAAACACACTTGGGAGATGACATATTACCTGATCAATCAAGCCACTG 1628
 QY 901 AAATTCAGCTTCAAGCAGAACCCAGACCTTCCCAAGCTGAGATTTCCAAAAGTT 960
 DB 1627 AAATTCAGCTTCAAGCAGAACCCAGACCTTCCCAAGCTGAGATTTCCAAAAGTT 1568
 QY 961 GAGAAATGATGACTTCAAAAGCTGAAAGCTTCAGATTGAAGGATCAATCCGACATCA 1020
 DB 1567 GAGAAATGATGACTTCAAAAGCTGAAAGCTTCAGATTGAAGGATCAATCCGACATCA 1508
 QY 1021 ACTATTAAATGAAATGCTGTTTAGGGTCTTTCAAAGAGCTGAAAGATATTGTCA 1080
 DB 1507 ACTATTAAATGAAATGCTGTTTAGGGTCTTTCAAAGAGCTGAAAGATATTGTCA 1448
 QY 1081 GTCTTTAGGGGTTTGGCTGATGCGAGGAAAGTCTTTTGTCT 1126
 DB 1447 GTCTTTAGGGGTTTGGCTGATGCGAGGAAAGTCTTTTGTCT 1402

RESULT 15

US-09-724-676-34565/c
 ; Sequence 34565, Application US/09724676
 ; GENERAL INFORMATION:
 ; APPLICANT: Compugen LTD
 ; TITLE OF INVENTION: Variants of alternative splicing
 ; FILE REFERENCE: 129181.4 Compugen
 ; CURRENT APPLICATION NUMBER: US/09/724,676
 ; CURRENT FILING DATE: 2000-11-28
 ; NUMBER OF SEQ ID NOS: 97222
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 34565
 ; LENGTH: 3023
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-724-676-34565

Query Match 66.7%; Score 1024; DB 5; Length 3023;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1124; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGGGGGGGGGGGAGACCACTTGGCTGCTCCGCGCGCGCACTTGGCTGCTCCGT 60
 DB 3023 GGGGGGGGGGGGAGACCACTTGGCTGCTCCGCGCGCGCGCGCGCGCGCGCGCG 2964
 QY 61 CCGCGCGCGCACTTGGCTGCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
 DB 2963 CCGCGCGCGCACTTGGCTGCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2904
 QY 121 TCGAGCTGCG 180

Qy	301	AGCTGAGATGATTAATTCCTCTCTGTGTCAACAACAAGAGTGTCTGGAGAGGAGTTTG	360
Db	2191	AGCTGAGAGGTGAATTCCTCTCTGTGTCAACAACAAGAGTGTCTGGAGAGGAGTTTG	2132
Qy	361	GAGAGGTTGCTGTGGTTTATCAAGGSAATCAAAATGCTAAAGAGTGTCTCAAGGSA	420
Db	2131	GAGAGGTTGCTGTGGTTTATCAAGGSAATCAAAATGCTAAAGAGTGTCTCAAGGSA	2072
Qy	421	GAGAAATCTGGGATCCCAAGATCCCAAGATCTTTTGGACAGCTGTGGATTCCTCAC	480
Db	2071	GAGAAATCTGGGATCCCAAGATCCCAAGATCTTTTGGACAGCTGTGGATTCCTCAC	2012
Qy	481	AGAGAAGAGGGGACTTGGGCCAGTTTATGGCTTCAATGAGAGCATTTTGGGGAGAA	540
Db	2011	AGAGAAGAGGGGACTTGGGCCAGTTTATGGCTTCAATGAGAGCATTTTGGGGAGAA	1952
Qy	541	TACAGATATGSAATCAGATTATTCAAGACAGGAGTTTGAACAATGSAAGATGATT	600
Db	1951	TACAGATATGSAATCAGATTATTCAAGACAGGAGTTTGAACAATGSAAGATGATT	1892
Qy	601	GACACATTCMAAACCAACCTCGAGACAGAAAGAAATCAATCATGAGGCTTGGAAATCA	660
Db	1891	GACACATTCMAAACCAACCTCGAGACAGAAAGAAATCAATCATGAGGCTTGGAAATCA	1832
Qy	661	GATCTTCCTCGATGGGCGGTGCTCCATGCAATGCCCTTCGACAGTCTATGTGTAA	720
Db	1831	GATCTTCCTCGATGGGCGGTGCTCCATGCAATGCCCTTCGACAGTCTATGTGTAA	1772
Qy	721	AGTGAAGTGTCCGACAGCTGTACAGAGATGAGGAGACATGGGCTGTGGTGTCTTTC	780
Db	1771	AGTGAAGTGTGTCCGACAGCTGTACAGAGATGAGGAGACATGGGCTGTGGTGTCTTTC	1712
Qy	781	AACATGCGACGCTACGCTGTGTACAGATGATTTGGGACATCAAGGGCTGTAAACCA	840
Db	1711	AACATGCGACGCTACGCTGTGTACAGATGATTTGGGACATCAAGGGCTGTAAACCA	1652
Qy	841	GGTGACTTATACACCTTTGGAGATGCAATATTACCTGAATCAACATGAGCGACATG	900
Db	1651	GGTGACTTATACACCTTTGGAGATGCAATATTACCTGAATCAACATGAGCGACATG	1592
Qy	901	AAAAATCAGCTCAGAGGAGAACCCAGACCTTCCCAAGGCTCAGATTCCTGAAAGTT	960
Db	1591	AAAAATCAGCTCAGAGGAGAACCCAGACCTTCCCAAGGCTCAGATTCCTGAAAGTT	1532
Qy	961	GAGAAATTTGATGACTTCAAGCTGAGACATTTAGATTGAGAGGTACATCCGATCA	1020
Db	1531	GAGAAATTTGATGACTTCAAGCTGAGACATTTAGATTGAGAGGTACATCCGATCA	1472
Qy	1021	ACTATTAAATGGAATGCTGTGTTTACAGTGTCTTCAAGAGCTGGAAGATTTGTCA	1080
Db	1471	ACTATTAAATGGAATGCTGTGTTTACAGTGTCTTCAAGAGCTGGAAGATTTGTCA	1412
Qy	1081	GTCCTTAGGGGTGGGCTGATGCCAGAGGTAAAGTCTTTTGTCT	1126
Db	1411	GTCCTTAGGGGTGGGCTGATGCCAGAGGTAAAGTCTTTTGTCT	1366
RESULT 13			
US-09-724-676-34548/c			
Sequence 34548, Application US/09724676			
GENERAL INFORMATION:			
APPLICANT: Comugen LTD			
TITLE OF INVENTION: Variants of alternative splicing			
FILE REFERENCE: 129181.4 Comugen			
CURRENT APPLICATION NUMBER: US/09/724, 676			
CURRENT FILING DATE: 2000-11-28			
NUMBER OF SEQ ID NOS: 97222			
SOFTWARE: PatentIn version 3.2			
SEQ ID NO 34548			
LENGTH: 2527			
TYPE: DNA			
ORGANISM: Homo sapiens			
FEATURE:			

```

; NAME/KEY: misc_feature
; LOCATION: (41).(41)
; OTHER INFORMATION: n is a,c,g, or t
US-09-724-676-34548

```

Query Match	66.7%;	Score 1024;	DB 5;	Length 2527;
Best Local Similarity	99.8%;	Pred. No. 0;		
Matches 112;	Conservative	0;	Mismatches 2;	Indels 0;
				Gaps 0;

[illegible]

Db 2051 GAGAGTTGCTGTTGTTTATCAAGGATCCAAATGCTTAAGAGCTGTTCCAAAGGA 1992
QY 421 GTGAAATCTGGATGCGAATGATCCGAGACTTTTGGACAAGCTGGATTTCCACC 480
Db 1991 GTGAAATCTGGATGCGAATGATCCGAGACTTTTGGACAAGCTGGATTTCCACC 1932
QY 481 AGAGAAGAGGGGACTTGGGGCCAGTTTATGCTTCCAGTGGAGGATTTTGGGGGAGA 540
Db 1931 AGAGAAGAGGGGACTTGGGGCCAGTTTATGCTTCCAGTGGAGGATTTTGGGGGAGA 1872
QY 541 TACAGAGATATGATATCAATTTATTCAGGACAGGAGATTTTACCAATGCAAAAGTAT 600
Db 1871 TACAGAGATATGATATCAATTTATTCAGGACAGGAGATTTTACCAATGCAAAAGTAT 1812
QY 601 GACACCATCAAAACCAACCCCTGACGACAGAGATCAATGTCAGTCTTGGATCCAGA 660
Db 1811 GACACCATCAAAACCAACCCCTGACGACAGAGATCAATGTCAGTCTTGGATCCAGA 1752
QY 661 GATCTTCTGATGAGGCTGCTGCTCCATGCTCCATGCTCCATGCTTCTTATGTTGATC 720
Db 1751 GATCTTCTGATGAGGCTGCTGCTCCATGCTCCATGCTCCATGCTTCTTATGTTGATC 1692
QY 721 AGTAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db 1691 AGTAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1632
QY 781 AACATGCGACGCTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Db 1631 AACATGCGACGCTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1572
QY 841 GGTGACTTTATACACCTTTGGAGATGCAATTTTACCTGATATCAATCCAGGACTG 900
Db 1571 GGTGACTTTATACACCTTTGGAGATGCAATTTTACCTGATATCAATCCAGGACTG 1512
QY 901 AAAATTCAGCTTCAGGAGAACCCAGACCTTTCCAAAGCTTCAGATTTTGGAAAGTT 960
Db 1511 AAAATTCAGCTTCAGGAGAACCCAGACCTTTCCAAAGCTTCAGATTTTGGAAAGTT 1452
QY 961 GAGAAATATGATGCTTCAAAAGCTGAAACCTTTCAATGATGAGGATCAATCCGATCA 1020
Db 1451 GAGAAATATGATGCTTCAAAAGCTGAAACCTTTCAATGATGAGGATCAATCCGATCA 1392
QY 1021 ACTATTTAAATGGAATGCTGTTTGGGCTTTCAAAGGACTGAAAGGATTTGCA 1080
Db 1391 ACTATTTAAATGGAATGCTGTTTGGGCTTTCAAAGGACTGAAAGGATTTGCA 1332
QY 1081 GTCTTAGGGGTTGGGCTGGATCCGAGGTAAAGTTCTTTTGGCT 1126
Db 1331 GTCTTAGGGGTTGGGCTGGATCCGAGGTAAAGTTCTTTTGGCT 1286

RESULT 10

US-09-724-676A-34535/c
; Sequence 34535, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Comugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Comugen
; CURRENT APPLICATION NUMBER: US/09/724.676A
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 34535
; LENGTH: 2411
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (41)..(41)
; OTHER INFORMATION: n is a,c,g, or t
US-09-724-676A-34535

Query Match

66.7%; Score 1024; DB 5; Length 2411;

Best Local Similarity 99.8%; Pred. No. 0;
Matches 1124; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 GGGGGGGGGGGGACCAATGAGCTGCTCCGCTCCGCGCGCACTTGGCTGCTCCGT 60
Db 2411 GGGGGGGGGGGGACCAATGAGCTGCTCCGCTCCGCGCGCACTTGGCTGCTCCGT 2352
QY 61 CCGCGCGCGCACTTGGCTGCTCCGCTCCGCGCGCGCGCGCGCGCGCGCGCGCG 120
Db 2351 CCGCGCGCGCACTTGGCTGCTCCGCTCCGCGCGCGCGCGCGCGCGCGCGCGCG 2292
QY 121 TCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
Db 2291 TCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2232
QY 181 CCGCGCGCGCGCACTTGGCTGCTCCGCTCCGCGCGCGCGCGCGCGCGCGCGCG 240
Db 2231 CCGCGCGCGCGCACTTGGCTGCTCCGCTCCGCGCGCGCGCGCGCGCGCGCGCG 2172
QY 241 AGAGAAGACACCGCAACCGGCGCAACCGGCGCAACCGGCGCAACCGGCGCAAC 300
Db 2171 AGAGAAGACACCGCAACCGGCGCAACCGGCGCAACCGGCGCAACCGGCGCAAC 2112
QY 301 AGCGTGAAGATGATTTCCCTGCTGCAACAACGATGCTTGAAGAGTGTGTTG 360
Db 2111 AGCGTGAAGATGATTTCCCTGCTGCAACAACGATGCTTGAAGAGTGTGTTG 2052
QY 361 GAGAGTGTGCTGCTGCTTATCAAGGATCCAAATGCTTAAAGCTGCTTCAAGGA 420
Db 2051 GAGAGTGTGCTGCTGCTTATCAAGGATCCAAATGCTTAAAGCTGCTTCAAGGA 1992
QY 421 GTGAAATCTGGATGCGAATGATCCGAGACTTTTGGACAGCTGGGATTTCCACC 480
Db 1991 GTGAAATCTGGATGCGAATGATCCGAGACTTTTGGACAGCTGGGATTTCCACC 1932
QY 481 AGAGAAGAGGGGACTTGGGGCCAGTTTATGCTTCCAGTGGAGGATTTTGGGGGAGA 540
Db 1931 AGAGAAGAGGGGACTTGGGGCCAGTTTATGCTTCCAGTGGAGGATTTTGGGGGAGA 1872
QY 541 TACAGAGATATGATATCAATTTATTCAGGACAGGAGATTTTACCAATGCAAAAGTAT 600
Db 1871 TACAGAGATATGATATCAATTTATTCAGGACAGGAGATTTTACCAATGCAAAAGTAT 1812
QY 601 GACACCATCAAAACCAACCCCTGACGACAGAGATCAATGTCAGTCTTGGATCCAGA 660
Db 1811 GACACCATCAAAACCAACCCCTGACGACAGAGATCAATGTCAGTCTTGGATCCAGA 1752
QY 661 GATCTTCTGATGAGGCTGCTGCTCCATGCTCCATGCTCCATGCTTCTTATGTTGATC 720
Db 1751 GATCTTCTGATGAGGCTGCTGCTCCATGCTCCATGCTCCATGCTTCTTATGTTGATC 1692
QY 721 AGTAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db 1691 AGTAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1632
QY 781 AACATGCGACGCTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Db 1631 AACATGCGACGCTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1572
QY 841 GGTGACTTTATACACCTTTGGAGATGCAATTTTACCTGATATCAATCCAGGACTG 900
Db 1571 GGTGACTTTATACACCTTTGGAGATGCAATTTTACCTGATATCAATCCAGGACTG 1512
QY 901 AAAATTCAGCTTCAGGAGAACCCAGACCTTTCCAAAGCTTCAGATTTTGGAAAGTT 960
Db 1511 AAAATTCAGCTTCAGGAGAACCCAGACCTTTCCAAAGCTTCAGATTTTGGAAAGTT 1452
QY 961 GAGAAATATGATGCTTCAAAAGCTGAAACCTTTCAATGATGAGGATCAATCCGATCA 1020
Db 1451 GAGAAATATGATGCTTCAAAAGCTGAAACCTTTCAATGATGAGGATCAATCCGATCA 1392
QY 1021 ACTATTTAAATGGAATGCTGTTTGGGCTTTCAAAGGACTGAAAGGATTTGCA 1080